OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 02:50:53 ; Search time 1614.64 Seconds
(without alignments)
722.185 Million cell updates/sec

Title: US-09-974-974-1
Perfect score: 32
Sequence: 1 gguccuggccugaugagagugaugagcucuuc 32

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 238946

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.8	49.4	55	9	BX655160	BX655160 Arabidops
	2	15.4	48.1	43	1	AA947665	AA947665 oq32d02.s
c	3	15.2	47.5	48	1	AA861970	AA861970 oi45d12.s
	4	15.2	47.5	57	9	DR1M24T	AL742081 Danio rer
c	5	15.2	47.5	59	8	BZ770511	BZ770511 SALK_1434
c	6	15	46.9	58	8	BH770391	BH770391 LLMGtag16
c	7	14.8	46.2	57	9	CR110068	CR110068 Forward s
	8	14.8	46.2	60	6	CD948943	CD948943 SAH_46 Ge
	9	14.6	45.6	32	8	AZ650179	AZ650179 1M0520P11
c	10	14.6	45.6	42	7	CF920754	CF920754 gmrhRww3-
	11	14.6	45.6	46	1	AV833578	AV833578 AV833578
c	12	14.6	45.6	52	2	AW246952	AW246952 2822591.5
	13	14.6	45.6	52	9	CR144482	CR144482 Forward s
	14	14.6	45.6	57	8	AZ819538	AZ819538 2M0091A14
	15	14.4	45.0	46	1	AA181388	AA181388 zp42f11.s
	16	14.4	45.0	46	9	CR398085	CR398085 Arabidops
	17	14.4	45.0	51	2	BF308407	BF308407 601890722
	18	14.4	45.0	52	1	AA173932	AA173932 zp01e02.s
c	19	14.4	45.0	54	2	AW781356	AW781356 sk68g12.y
c	20	14.4	45.0	55	2	BE748176	BE748176 601571564
c	21	14.2	44.4	42	1	AJ789426	AJ789426 AJ789426
	22	14.2	44.4	50	1	AU107710	AU107710 AU107710
c	23	14.2	44.4	52	2	BE321553	BE321553 NF024G09I
c	24	14.2	44.4	54	7	CK348290	CK348290 LeCfAFLP0
	25	14.2	44.4	58	9	AL753395	AL753395 Arabidops
c	26	14	43.8	39	7	R42069	R42069 yf98c06.s1
c	27	14	43.8	43	1	AI647472	AI647472 uk42e12.x
c	28	14	43.8	53	1	AA846708	AA846708 aj41b12.s
	29	14	43.8	53	6	CA339521	CA339521 NISC_1y03
	30	14	43.8	56	8	BH911451	BH911451 SALK_0687
c	31	14	43.8	58	8	AZ792055	AZ792055 2M0043E15
	32	14	43.8	58	9	CG723363	CG723363 1119076A0
	33	13.8	43.1	27	8	AZ993846	AZ993846 2M0279K05
c	34	13.8	43.1	49	1	AI869979	AI869979 wl63d04.x
	35	13.8	43.1	55	4	BG759188	BG759188 602710848
	36	13.8	43.1	55	8	AZ918644	AZ918644 1006004H0
c	37	13.8	43.1	58	1	AI039811	AI039811 ox49d04.x
	38	13.8	43.1	59	1	AI988246	AI988246 sc98a02.y
c	39	13.8	43.1	59	5	BP134619	BP134619 BP134619
c	40	13.8	43.1	59	7	CN933300	CN933300 000510AOF
c	41	13.6	42.5	40	6	CA586986	CA586986 LBG26p15
	42	13.6	42.5	45	7	T17565	T17565 mps v253 Th
	43	13.6	42.5	49	1	AA691190	AA691190 vt34c05.r
	44	13.6	42.5	52	1	AA674719	AA674719 vm74c06.s
c	45	13.6	42.5	54	8	BZ289822	BZ289822 SALK_0232

ALIGNMENTS

RESULT 1
BX655160/c
LOCUS BX655160 55 bp DNA linear GSS 04-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-587C03-021291, genomic survey sequence.
ACCESSION BX655160
VERSION BX655160.1 GI:37611548
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and Weisshaar,B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and Weisshaar,B.
TITLE High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
JOURNAL BioTechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4 (bases 1 to 55)
AUTHORS Li,Y., Strizhov,N., Rosso,M.G. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by BAC clone K18L3. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program

designated 'GABI'. Information on line availability can be found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.

```
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-587C03-021291"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pGABI1 (GenBank accession number: AY529716). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             49.4%;  Score 15.8;  DB 9;  Length 55;
  Best Local Similarity   51.9%;  Pred. No. 6.5e+04;
  Matches   14;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UGGCCUGAUGAGAGUGAUGAGCUCUUC 32
              : | |:||:||| | || ||  :|::|
Db         47 TCGACTGATGAGCGAGAAGAATTCTTC 21


RESULT 2
AA947665
LOCUS       AA947665                 43 bp    mRNA    linear   EST 23-JUL-1998
DEFINITION  oq32d02.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1588035 3'
            similar to TR:Q13539 Q13539 MARINER TRANSPOSASE. ;, mRNA sequence.
ACCESSION   AA947665
VERSION     AA947665.1  GI:3108918
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1878   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1588035"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GC4"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             48.1%;  Score 15.4;  DB 1;  Length 43;
  Best Local Similarity   76.5%;  Pred. No. 9.2e+04;
  Matches   13;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 UGAUGAGAGUGAUGAGC 27
              :||:||||| ||:||||
Db          3 TGATGAGAGGGATGAGC 19


RESULT 3
AA861970/c
LOCUS       AA861970                 48 bp    mRNA    linear   EST 13-APR-1999
DEFINITION  oi45d12.s1 NCI_CGAP_HN3 Homo sapiens cDNA clone IMAGE:1485623 3'
            similar to gb:Z19574_rna1 KERATIN, TYPE I CYTOSKELETAL 17 (HUMAN);,
            mRNA sequence.
ACCESSION   AA861970
VERSION     AA861970.1  GI:2954449
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: John Ensley, M.D.
             cDNA Library Preparation: Stratagene, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1607   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1485623"
                     /tissue_type="squamous cell carcinoma from base of tongue"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="NCI_CGAP_HN3"
                     /note="Organ: tongue; Vector: Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Average insert size 1.0 kb. 5' adaptor sequence:
                     5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                     (GA)10ACTAGTCTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             47.5%;  Score 15.2;  DB 1;  Length 48;
  Best Local Similarity   46.4%;  Pred. No. 1.1e+05;
  Matches   13;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 CCUGGCCUGAUGAGAGUGAUGAGCUCUU 31
              ||: |||:||:|| | : | |  |: ::
Db         36 CCTTGCCTGATGACAATAAAGCTCTGTT 9
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 01:32:47 ; Search time 202.305 Seconds
(without alignments)
830.337 Million cell updates/sec

Title: US-09-974-974-1
Perfect score: 32
Sequence: 1 gguccuggccugaugagagugaugagcucuuc 32

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 4135388

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	32	100.0	32	6	AAL40446	Aal40446 Maxizyme-
	2	21.4	66.9	33	6	AAL40448	Aal40448 Maxizyme-
	3	17.8	55.6	60	6	ABN44085	Abn44085 Human spl
	4	17.4	54.4	29	2	AAA16986	Aaa16986 Aryl hydr
	5	17.4	54.4	40	6	AAL40455	Aal40455 Maxizyme
	6	16.6	51.9	30	2	AAZ32140	Aaz32140 Human PRO
	7	16.6	51.9	30	3	AAZ89576	Aaz89576 Human PRO
	8	16.6	51.9	30	3	AAZ89588	Aaz89588 Human PRO
	9	16.6	51.9	30	3	AAA46937	Aaa46937 Probe use
	10	16.6	51.9	30	10	ADJ58635	Adj58635 Human PRO
c	11	16.4	51.2	27	6	ABS66252	Abs66252 Anti-huma
	12	16.4	51.2	45	2	AAT96613	Aat96613 Cyclitol-
	13	16.4	51.2	55	4	AAS14065	Aas14065 Vector pA
	14	16.2	50.6	39	2	AAV54276	Aav54276 Bacteriop
	15	16.2	50.6	39	2	AAX55358	Aax55358 Soluble s
c	16	16.2	50.6	47	3	AAZ69149	Aaz69149 Human map
c	17	16.2	50.6	51	5	AAF82327	Aaf82327 RAR-gamma
c	18	16.2	50.6	60	6	ABN47364	Abn47364 Human spl
	19	16	50.0	29	2	AAA19290	Aaa19290 Integrin
	20	16	50.0	34	10	ADD28468	Add28468 Escherich
c	21	16	50.0	34	10	ADD28469	Add28469 Escherich
	22	16	50.0	53	2	AAQ40286	Aaq40286 Sequence
	23	16	50.0	53	3	AAA89859	Aaa89859 Primer P-
	24	16	50.0	53	4	AAS12777	Aas12777 PCR prime
c	25	15.8	49.4	38	4	ABK04167	Abk04167 Human NOG
c	26	15.8	49.4	38	11	ADL56413	Adl56413 Human PKR
c	27	15.6	48.8	48	9	ADA52426	Ada52426 Human gra
	28	15.6	48.8	50	4	AAI77405	Aai77405 Human sil
	29	15.6	48.8	60	6	ABN50812	Abn50812 Human spl
c	30	15.4	48.1	25	6	ABV92265	Abv92265 Human POS
c	31	15.4	48.1	25	6	ABV92268	Abv92268 Human POS
	32	15.4	48.1	36	2	AAT56380	Aat56380 Mouse TNF
	33	15.4	48.1	36	6	ABX01861	Abx01861 HCV hamme
c	34	15.4	48.1	47	2	AAV57986	Aav57986 Murine PS
c	35	15.4	48.1	47	3	AAA07303	Aaa07303 PCR prime
c	36	15.4	48.1	51	4	AAH90587	Aah90587 Human clo
	37	15.4	48.1	60	6	ABN32308	Abn32308 Human spl
c	38	15.2	47.5	20	12	ADJ85725	Adj85725 Nucleic a
c	39	15.2	47.5	34	3	AAZ59080	Aaz59080 Primer SK
c	40	15.2	47.5	50	6	ABZ04267	Abz04267 Human leu
c	41	15.2	47.5	60	6	ABN48062	Abn48062 Human spl
c	42	15	46.9	25	6	ABV92264	Abv92264 Human POS
c	43	15	46.9	25	6	ABV92263	Abv92263 Human POS
	44	15	46.9	27	3	AAZ61898	Aaz61898 Hammerhea
	45	15	46.9	29	2	AAA22030	Aaa22030 Integrin

ALIGNMENTS

RESULT 1
AAL40446
ID AAL40446 standard; RNA; 32 BP.
XX
AC AAL40446;
XX
DT 19-SEP-2002 (first entry)
XX
DE Maxizyme-constituting RNA molecule T-MzR.
XX
KW Enzyme; modifiable RNA cleavage activity; maxizyme-constituting RNA;
KW trans maxizyme; T-MzR; ss.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT misc_binding 1. .9
FT /*tag= a
FT /bound_moiety= "Bcl-2_mRNA"
FT /note= "Forms a double-stranded region with nucleotides
FT 19-11 of sequence AAL40451"
FT misc_binding 11. .19
FT /*tag= b
FT /bound_moiety= "T-MzL_RNA"
FT /note= "Forms a double-stranded region with nucleotides
FT 16-8 of sequence AAL40447"
FT misc_binding 14. .19
FT /*tag= e
FT /bound_moiety= "T-MzL_RNA"
FT /note= "Forms a double-stranded region with nucleotides
FT 18-13 of sequence AAL40447"
FT misc_binding 20. .32
FT /*tag= f
FT /bound_moiety= "HIV-1_tat_mRNA"
FT /note= "Forms a double-stranded region with nucleotides
FT 13-1 of sequence AAL40452"
FT misc_binding 21. .22
FT /*tag= c
FT /bound_moiety= "T-MzL_RNA"
FT /note= "Forms a double-stranded region with nucleotides 6
FT -5 of sequence AAL40447"
FT misc_binding 24. .25
FT /*tag= d
FT /bound_moiety= "T-MzL_RNA"

```
FT                  /note= "Forms a double-stranded region with nucleotides 3
FT                  -2 of sequence AAL40447"
XX
PN   JP2002119283-A.
XX
PD   23-APR-2002.
XX
PF   13-OCT-2000; 2000JP-00313320.
XX
PR   13-OCT-2000; 2000JP-00313320.
XX
PA   (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.
XX
DR   WPI; 2002-483792/52.
XX
PT   A nucleic acid enzyme which has selective and effective eradicating
PT   activity towards harmful cells by acquiring cleavage activity of a
PT   specific target RNA by recognition of the other RNA molecule.
XX
PS   Claim 2; Page 2; 17pp; Japanese.
XX
CC   The invention relates to a nucleic acid enzyme with modifiable RNA
CC   cleavage activity. More specifically the invention relates to a nucleic
CC   acid enzyme, trans maxizyme, which has selective and effective
CC   eradicating activity towards harmful cells by acquiring cleavage activity
CC   of a specific target RNA by recognition of the other RNA molecule. The
CC   enzyme of the invention is useful for cleaving target RNA and is useful
CC   in treating diseases caused by the target RNA. This polynucleotide
CC   sequence represents the maxizyme-constituting RNA molecule T-MzR relating
CC   to the invention
XX
SQ   Sequence 32 BP; 5 A; 7 C; 11 G; 0 T; 9 U; 0 Other;

  Query Match             100.0%;  Score 32;  DB 6;  Length 32;
  Best Local Similarity   100.0%;  Pred. No. 0.00034;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAUGAGCUCUUC 32
              ||||||||||||||||||||||||||||||||
Db          1 GGUCCUGGCCUGAUGAGAGUGAUGAGCUCUUC 32


RESULT 2
AAL40448
ID   AAL40448 standard; RNA; 33 BP.
XX
AC   AAL40448;
XX
DT   19-SEP-2002  (first entry)
XX
DE   Maxizyme-constituting RNA molecule B-MzR.
XX
KW   Enzyme; modifiable RNA cleavage activity; maxizyme-constituting RNA;
KW   trans maxizyme; B-MzR; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   misc_binding    1. .9
FT                   /*tag=  a
FT                   /bound_moiety= "Bcl-2_mRNA"
FT                   /note= "Forms a double-stranded region with nucleotides
FT                   19-11 of sequence AAL40451"
FT   misc_binding    12. .19
FT                   /*tag=  d
FT                   /bound_moiety= "b-MzL_RNA"
FT                   /note= "Forms a double-stranded region with nucleotides
FT                   17-10 of sequence AAL40449"
FT   misc_binding    14. .19
FT                   /*tag=  b
FT                   /bound_moiety= "B-MzL_RNA"
FT                   /note= "Forms a double-stranded region with nucleotides
FT                   20-15 of sequence AAL40449"
FT   misc_binding    20. .33
FT                   /*tag=  c
FT                   /bound_moiety= "BCR-ABL_mRNA"
FT                   /note= "Forms a double-stranded region with nucleotides
FT                   14-1 of sequence AAL40453"
XX
PN   JP2002119283-A.
XX
PD   23-APR-2002.
XX
PF   13-OCT-2000; 2000JP-00313320.
XX
PR   13-OCT-2000; 2000JP-00313320.
XX
PA   (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.
XX
DR   WPI; 2002-483792/52.
XX
PT   A nucleic acid enzyme which has selective and effective eradicating
PT   activity towards harmful cells by acquiring cleavage activity of a
PT   specific target RNA by recognition of the other RNA molecule.
XX
PS   Claim 2; Page 2; 17pp; Japanese.
XX
CC   The invention relates to a nucleic acid enzyme with modifiable RNA
CC   cleavage activity. More specifically the invention relates to a nucleic
CC   acid enzyme, trans maxizyme, which has selective and effective
CC   eradicating activity towards harmful cells by acquiring cleavage activity
CC   of a specific target RNA by recognition of the other RNA molecule. The
CC   enzyme of the invention is useful for cleaving target RNA and is useful
CC   in treating diseases caused by the target RNA. This polynucleotide
CC   sequence represents the maxizyme-constituting RNA molecule B-MzR relating
CC   to the invention
XX
SQ   Sequence 33 BP; 6 A; 5 C; 12 G; 0 T; 10 U; 0 Other;

  Query Match             66.9%;  Score 21.4;  DB 6;  Length 33;
  Best Local Similarity   95.7%;  Pred. No. 14;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAU 23
              |||||||||||||||||||| ||
Db          1 GGUCCUGGCCUGAUGAGAGUUAU 23


RESULT 3
ABN44085
ID   ABN44085 standard; DNA; 60 BP.
XX
AC   ABN44085;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:16833.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB001903.
XX
PR   28-JUL-2000; 2000US-0221607P.
PR   02-MAY-2001; 2001US-0287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
```

```
; FastDB parameter file
; Created by tport on Mon 1 Nov 104 16:34:33-PST.

Query-File-Name                     US09111292.seq
Query-Sequence-Name-or-Number       US-09-111-292-6
Query-Sequence-Range                1 1099
File-to-Search                      US09111292.seq
Search-Name                         us-09-111-292-6
Similarity-matrix                   Unitary
Number-of-PAMs                      150
Threshold-level-of-similarity       100
Ktuple                              4
Mismatch-penalty                    1
Joining-penalty                     30
Cutoff-score                        1
Randomization-group-length          0
Number-of-randomizations            0
Translation-frame                   0
Window-size                         500
Gap-penalty                         5.00
Gap-size-penalty                    0.33
Time-to-submit-batch-job            none
Number-of-initial-scores-to-save    2
Number-of-optimized-scores-to-save  20
Number-of-alignments-to-save        2
Display-annotations                 No
Display-context                     50
Notify-upon-completion              Yes
Number-of-sequences-to-collect      13
Number-to-save-in-indirect-file     13
```

```
Qy          61 UUCGAAACCGGGCACUACAAAAACCAAC 88
               ::|||||||||||||:|||||||||||||
Db          61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 12
AX138489
LOCUS       AX138489                142 bp    RNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 50 from Patent EP1097993.
ACCESSION   AX138489
VERSION     AX138489.1  GI:14274385
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
  TITLE     Functional ribozyme chimeric molecules capable of sliding
  JOURNAL   Patent: EP 1097993-A 50 09-MAY-2001;
            Secretary of Agency of Industrial Science and Technology (JP) ;
            Taira, Kazunari (JP)
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="the nucleotide sequence of CPP Rz2"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 142;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 13
AX138490
LOCUS       AX138490                142 bp    RNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 51 from Patent EP1097993.
ACCESSION   AX138490
VERSION     AX138490.1  GI:14274386
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
  TITLE     Functional ribozyme chimeric molecules capable of sliding
  JOURNAL   Patent: EP 1097993-A 51 09-MAY-2001;
            Secretary of Agency of Industrial Science and Technology (JP) ;
            Taira, Kazunari (JP)
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="the nucleotide sequence of CPP Rz3"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 142;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 14
AX138492
LOCUS       AX138492                142 bp    RNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 53 from Patent EP1097993.
ACCESSION   AX138492
VERSION     AX138492.1  GI:14274388
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
  TITLE     Functional ribozyme chimeric molecules capable of sliding
  JOURNAL   Patent: EP 1097993-A 53 09-MAY-2001;
            Secretary of Agency of Industrial Science and Technology (JP) ;
            Taira, Kazunari (JP)
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="the nucleotide sequence of CPP Rz5"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 142;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 15
BD015656
LOCUS       BD015656                142 bp    RNA     linear   PAT 27-AUG-2002
DEFINITION  Slidable functional chimeric molecule.
ACCESSION   BD015656
VERSION     BD015656.1  GI:22556793
KEYWORDS    JP 2001190282-A/49.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 142)
  AUTHORS   Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
  TITLE     Slidable functional chimeric molecule
  JOURNAL   Patent: JP 2001190282-A 49 17-JUL-2001;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND MANABU NAKAJIMA TECHNOLOGY, KAZUNARI TAIRA
COMMENT     OS   Artificial Sequence
            PN   JP 2001190282-A/49
            PD   17-JUL-2001
            PF   02-NOV-2000 JP 2000336082
            PI   KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO KUWABARA,HIROAKI  PI
            KAWASAKI
            PC
            C12N15/09,A61K31/7105,A61K31/711,A61K38/00,A61K48/00,A61P31/12, PC
             A61P43/00,
            PC   C12N9/22,C12Q1/02,C12Q1/68,G01N33/53,G01N33/566,C12N15/00, PC
             A61K37/02
            CC   Description of Artificial Sequence: the
            nucleotide sequence of
```

```
            CC              CPP Rz1
            FH   Key             Location/Qualifiers
            FT   source          1. .142
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 142;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


Search completed: October 31, 2004, 04:50:41
Job time : 1507 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 03:24:30 ; Search time 1505 Seconds
(without alignments)
2765.112 Million cell updates/sec

Title: US-09-974-974-5
Perfect score: 88
Sequence: 1 accguugguuuccguagugu.....cgggcacuacaaaaaccaac 88

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 2719732

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	88	100.0	88	6	BD182356	BD182356 Novel max
	2	88	100.0	88	6	AX453846	AX453846 Sequence
	3	88	100.0	88	6	BD143502	BD143502 Nucleic a
	4	88	100.0	100	6	E47174	E47174 Method for
	5	88	100.0	117	6	BD174675	BD174675 Ribozyme
	6	88	100.0	137	6	AX429079	AX429079 Sequence
	7	88	100.0	137	6	BD143601	BD143601 Method of
	8	88	100.0	138	6	AX453858	AX453858 Sequence
	9	88	100.0	141	6	AX138491	AX138491 Sequence
	10	88	100.0	141	6	BD015659	BD015659 Slidable
	11	88	100.0	142	6	AX138488	AX138488 Sequence
	12	88	100.0	142	6	AX138489	AX138489 Sequence
	13	88	100.0	142	6	AX138490	AX138490 Sequence
	14	88	100.0	142	6	AX138492	AX138492 Sequence
	15	88	100.0	142	6	BD015656	BD015656 Slidable
	16	88	100.0	142	6	BD015657	BD015657 Slidable
	17	88	100.0	142	6	BD015658	BD015658 Slidable
	18	88	100.0	142	6	BD015660	BD015660 Slidable
	19	88	100.0	151	6	AX138449	AX138449 Sequence
	20	88	100.0	151	6	BD015617	BD015617 Slidable
	21	88	100.0	175	6	BD174679	BD174679 Ribozyme
	22	88	100.0	175	6	BD174680	BD174680 Ribozyme
	23	82.8	94.1	136	6	E33203	E33203 Expression
	24	82.8	94.1	153	12	AB080619	AB080619 Synthetic
	25	82.8	94.1	153	12	AB080620	AB080620 Synthetic
	26	82.8	94.1	153	12	AB080621	AB080621 Synthetic
	27	82.8	94.1	153	12	AB080622	AB080622 Synthetic
	28	82.8	94.1	153	12	AB080623	AB080623 Synthetic
	29	82.8	94.1	153	12	AB080624	AB080624 Synthetic
	30	81	92.0	142	6	E33204	E33204 Expression
	31	76	86.4	128	6	E33205	E33205 Expression
	32	74.8	85.0	95	6	E33206	E33206 Expression
	33	74.8	85.0	95	6	AX138447	AX138447 Sequence
	34	74.8	85.0	95	6	BD015615	BD015615 Slidable
c	35	74	84.1	183	11	G31283	G31283 sy905g1-L H
	36	73.4	83.4	149	6	E33207	E33207 Expression
	37	70.4	80.0	132	6	AR041184	AR041184 Sequence
	38	70.4	80.0	132	6	AR113038	AR113038 Sequence
	39	70.4	80.0	132	6	I66463	I66463 Sequence 17
c	40	66.8	75.9	113	6	E33209	E33209 Expression
	41	66	75.0	110	6	E33208	E33208 Expression
	42	65	73.9	75	10	RATTRV1	K00256 Rat liver V
	43	64.4	73.2	76	4	RABTRV1	K00253 Rabbit Val-
	44	64.4	73.2	76	9	HUMTRV1A	K00255 Human Val-t
	45	64.4	73.2	76	9	HUMTRV1B	K00254 Human Val-t

ALIGNMENTS

RESULT 1
BD182356
LOCUS BD182356 88 bp RNA linear PAT 15-MAY-2003
DEFINITION Novel maxizyme.
ACCESSION BD182356
VERSION BD182356.1 GI:30793274
KEYWORDS WO 02092821-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 88)
AUTHORS Taira,K., Warashina,T., Warashina,M., Kawasaki,H., Hara,T. and Nozawa,I.
TITLE Novel maxizyme
JOURNAL Patent: WO 02092821-A 3 21-NOV-2002;
NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY, TAISHO PHARMACEUTICAL CO LTD,GENOFUNCTION INC,KAZUNARI TAIRA, TOMOKO WARASHINA,MASAKI WARASHINA,HIROAKI KAWASAKI,TOSHIFUMI HARA, IWAO NOZAWA
COMMENT OS Homo sapiens (human)
PN WO 02092821-A/3
PD 21-NOV-2002
PF 30-APR-2002 WO 2002JP004322
PR 01-MAY-2001 JP 01P 134469
PI KAZUNARI TAIRA,TOMOKO WARASHINA,MASAKI WARASHINA,HIROAKI PI KAWASAKI,
PI TOSHIFUMI HARA,IWAO NOZAWA
PC C12N15/55,C12N9/22,C12Q1/34,A61K48/00,A61K38/43 CC Novel maxizyme
FH Key Location/Qualifiers
FT source 1. .87
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .88
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 88; DB 6; Length 88;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;

```
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 2
AX453846
LOCUS       AX453846                 88 bp    RNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 5 from Patent EP1213351.
ACCESSION   AX453846
VERSION     AX453846.1  GI:21713515
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzymes acquiring an activity for cleaving a target
            rna by recognising another molecule
  JOURNAL   Patent: EP 1213351-A 5 12-JUN-2002;
            National Institute of Advanced Industrial Science and Technology
            (JP)
FEATURES             Location/Qualifiers
     source          1. .88
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="tRNAVal promoter sequence"
ORIGIN

  Query Match           100.0%;  Score 88;  DB 6;  Length 88;
  Best Local Similarity  78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 3
BD143502
LOCUS       BD143502                 88 bp    RNA     linear   PAT 17-JAN-2003
DEFINITION  Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule.
ACCESSION   BD143502
VERSION     BD143502.1  GI:27849260
KEYWORDS    JP 2002119283-A/5.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule
  JOURNAL   Patent: JP 2002119283-A 5 23-APR-2002;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND P T BAKRIE AND BROTHERS,JAPAN BIO INDUSTRY ASSOCIATION,
            TECHNOLOGY
COMMENT     OS   Artificial Sequence
            PN   JP 2002119283-A/5
            PD   23-APR-2002
            PF   13-OCT-2000 JP 2000313320
            PI   KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO WARASHINA PC
            C12N15/09,A61K9/127,A61K38/46,A61K48/00,A61P31/12,A61P35/00, PC
            C12N1/15,
            PC   C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12Q1/25,C12Q1/68,C12N15/
            PC   00,A61K37/54,
            PC   C12N5/00
            CC   Description of Artificial Sequence: tRNAVal promoter sequence
            FH   Key             Location/Qualifiers
            FT   source          1. .88
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .88
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 88;  DB 6;  Length 88;
  Best Local Similarity  78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 4
E47174
LOCUS       E47174                   100 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Method for selecting active ribozyme.
ACCESSION   E47174
VERSION     E47174.1  GI:22553362
KEYWORDS    JP 2001128682-A/9.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Tabira,K. and Kawasaki,H.
  TITLE     Method for selecting active ribozyme
  JOURNAL   Patent: JP 2001128682-A 9 15-MAY-2001;
            AGENCY OF IND SCIENCE & TECHNOL
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001128682-A/9
            PD   15-MAY-2001
            PF   05-NOV-1999 JP 1999314579
            PI   KAZUNARI TABIRA,HIROAKI KAWASAKI
            PC   C12N15/09,C12Q1/68//C12N5/10,C12N15/00,C12N5/00 CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 88;  DB 6;  Length 100;
  Best Local Similarity  78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:|||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 5
```

```
BD174675
LOCUS       BD174675                117 bp    DNA     linear   PAT 18-MAR-2003
DEFINITION  Ribozyme expression system.
ACCESSION   BD174675
VERSION     BD174675.1  GI:29120365
KEYWORDS    JP 2002262880-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 117)
  AUTHORS   Takebe,Y. and Okawa,J.
  TITLE     Ribozyme expression system
  JOURNAL   Patent: JP 2002262880-A 1 17-SEP-2002;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF INFECTIONS DISEASES,
            YUTAKA TAKEBE,JUN OKAWA
COMMENT     OS   Artificial Sequence
            PN   JP 2002262880-A/1
            PD   17-SEP-2002
            PF   09-MAR-2001 JP 2001067253
            PI   YUTAKA TAKEBE,JUN OKAWA
            PC   C12N15/09,A61K31/711,A61K35/76,A61K48/00,A61P31/18,C12N5/10,
            PC   C12N9/00,
            PC   C12N15/00,C12N5/00
            CC   Description of Artificial Sequence: the nucleotide sequence of
            CC   tetO1-tRNAval
            FH   Key             Location/Qualifiers
            FT   source          1. .117
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 117;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::|||||:|||||||||||||:||||||
Db         29 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 88

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         89 TTCGAAACCGGGCACTACAAAAACCAAC 116


RESULT 6
AX429079
LOCUS       AX429079                137 bp    RNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 2 from Patent EP1201751.
ACCESSION   AX429079
VERSION     AX429079.1  GI:21540419
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K. and Sano,M.
  TITLE     Method for selecting highly functional nucleic acid molecules
            within cells
  JOURNAL   Patent: EP 1201751-A 2 02-MAY-2002;
            National Institute of Advanced Industrial Science and Technology
            (JP)
FEATURES             Location/Qualifiers
     source          1. .137
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="Sequence of tRNA-Luc GUA Rz"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 137;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::|||||:|||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 7
BD143601
LOCUS       BD143601                137 bp    RNA     linear   PAT 17-JAN-2003
DEFINITION  Method of selecting high-function nucleic acid molecule in cell.
ACCESSION   BD143601
VERSION     BD143601.1  GI:27849359
KEYWORDS    JP 2002125685-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 137)
  AUTHORS   Taira,K. and Sano,M.
  TITLE     Method of selecting high-function nucleic acid molecule in cell
  JOURNAL   Patent: JP 2002125685-A 2 08-MAY-2002;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND HIROSHI YATSUHASHI, TECHNOLOGY, KAZUNARI TAIRA
COMMENT     OS   Artificial Sequence
            PN   JP 2002125685-A/2
            PD   08-MAY-2002
            PF   30-OCT-2000 JP 2000331347
            PI   KAZUNARI TAIRA,MASAYUKI SANO
            PC   C12N15/09,C12N9/00,C12Q1/02,C12Q1/25,C12Q1/68,C12N15/00 CC
            Description of Artificial Sequence: Sequence of tRNA-Luc GUA  CC
                 Rz
            FH   Key             Location/Qualifiers
            FT   source          1. .137
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .137
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 137;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::|||||:|||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 8
AX453858
LOCUS       AX453858                138 bp    RNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 17 from Patent EP1213351.
ACCESSION   AX453858
VERSION     AX453858.1  GI:21713527
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzymes acquiring an activity for cleaving a target
```

```
          rna by recognising another molecule
  JOURNAL   Patent: EP 1213351-A 17 12-JUN-2002;
            National Institute of Advanced Industrial Science and Technology
            (JP)
FEATURES             Location/Qualifiers
     source          1. .138
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="tRNAVal T-MzL"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 138;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 9
AX138491
LOCUS       AX138491                 141 bp    RNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 52 from Patent EP1097993.
ACCESSION   AX138491
VERSION     AX138491.1  GI:14274387
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
  TITLE     Functional ribozyme chimeric molecules capable of sliding
  JOURNAL   Patent: EP 1097993-A 52 09-MAY-2001;
            Secretary of Agency of Industrial Science and Technology (JP) ;
            Taira, Kazunari (JP)
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="the nucleotide sequence of CPP Rz4"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 141;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 10
BD015659
LOCUS       BD015659                 141 bp    RNA     linear   PAT 27-AUG-2002
DEFINITION  Slidable functional chimeric molecule.
ACCESSION   BD015659
VERSION     BD015659.1  GI:22556796
KEYWORDS    JP 2001190282-A/52.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
  TITLE     Slidable functional chimeric molecule
  JOURNAL   Patent: JP 2001190282-A 52 17-JUL-2001;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND MANABU NAKAJIMA TECHNOLOGY, KAZUNARI TAIRA
COMMENT     OS   Artificial Sequence
            PN   JP 2001190282-A/52
            PD   17-JUL-2001
            PF   02-NOV-2000 JP 2000336082
            PI   KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO KUWABARA,HIROAKI  PI
            KAWASAKI
            PC
            C12N15/09,A61K31/7105,A61K31/711,A61K38/00,A61K48/00,A61P31/12, PC
             A61P43/00,
            PC   C12N9/22,C12Q1/02,C12Q1/68,G01N33/53,G01N33/566,C12N15/00, PC
             A61K37/02
            CC   Description of Artificial Sequence: the
            nucleotide sequence of
            CC            CPP Rz4
            FH   Key             Location/Qualifiers
            FT   source          1. .141
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 141;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60

Qy         61 UUCGAAACCGGGCACUACAAAAACCAAC 88
              ::|||||||||||||:|||||||||||||
Db         61 TTCGAAACCGGGCACTACAAAAACCAAC 88


RESULT 11
AX138488
LOCUS       AX138488                 142 bp    RNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 49 from Patent EP1097993.
ACCESSION   AX138488
VERSION     AX138488.1  GI:14274384
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
  TITLE     Functional ribozyme chimeric molecules capable of sliding
  JOURNAL   Patent: EP 1097993-A 49 09-MAY-2001;
            Secretary of Agency of Industrial Science and Technology (JP) ;
            Taira, Kazunari (JP)
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="the nucleotide sequence of CPP Rz1"
ORIGIN

  Query Match             100.0%;  Score 88;  DB 6;  Length 142;
  Best Local Similarity   78.4%;  Pred. No. 1.4e-19;
  Matches   69;  Conservative   19;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACCGUUGGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGG 60
              ||||::||:::|||:||:|:||:||::|:||||::||||:||||||||||||||:||||||
Db          1 ACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTAACACGCGAAAGGTCCCCGG 60
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 02:50:53 ; Search time 1362.36 Seconds
(without alignments)
722.185 Million cell updates/sec

Title: US-09-974-974-2
Perfect score: 27
Sequence: 1 gucugacuguucaucgaaaccgggucc 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 238946

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.4	53.3	50	1	AU102657	AU102657 AU102657
c	2	14.4	53.3	60	4	BI945126	BI945126 sb26b03.y
	3	14	51.9	51	1	AJ792078	AJ792078 AJ792078
	4	13.8	51.1	43	9	AL765834	AL765834 Arabidops
c	5	13.8	51.1	48	1	AV834227	AV834227 AV834227
	6	13.8	51.1	49	8	BH896046	BH896046 3526_1_36
c	7	13.8	51.1	52	9	AJ593160	AJ593160 Arabidops
	8	13.6	50.4	23	8	AZ475845	AZ475845 1M0294P09
c	9	13.4	49.6	60	9	CL314996	CL314996 mth2-133P
	10	13.2	48.9	43	9	CL265635	CL265635 02S0288-0
c	11	13.2	48.9	45	8	AZ820793	AZ820793 2M0093B18
	12	13.2	48.9	46	1	AI119755	AI119755 uc20a09.r
	13	13.2	48.9	53	9	TA126C12Q	AL463403 T. brucei
c	14	13	48.1	43	1	AA905188	AA905188 ok10b03.s
	15	13	48.1	44	8	BZ594864	BZ594864 SALK_0853
	16	13	48.1	54	9	AL752543	AL752543 Arabidops
	17	12.8	47.4	47	1	AA897045	AA897045 al08f07.s
c	18	12.8	47.4	51	8	BH789953	BH789953 SALK_0528
c	19	12.8	47.4	59	1	AU259987	AU259987 AU259987
	20	12.6	46.7	43	7	CO735543	CO735543 SlLL05b12
	21	12.6	46.7	45	7	D19978	D19978 HUMGS00943
	22	12.6	46.7	46	7	R86026	R86026 yp12f04.r1
	23	12.6	46.7	50	1	AU102358	AU102358 AU102358
	24	12.6	46.7	50	1	AU102359	AU102359 AU102359
	25	12.6	46.7	50	1	AU102360	AU102360 AU102360
	26	12.6	46.7	57	1	AA108111	AA108111 mm91c01.r
	27	12.6	46.7	58	4	BI903269	BI903269 id24c09.y
	28	12.6	46.7	58	4	BI903329	BI903329 id25b12.y
	29	12.6	46.7	58	4	BI966898	BI966898 id26a02.y
	30	12.6	46.7	58	5	BX624586	BX624586 BX624586
	31	12.6	46.7	58	7	D19112	D19112 MUSGS01323
	32	12.6	46.7	58	9	CR131887	CR131887 Forward s
c	33	12.4	45.9	31	8	BH791957	BH791957 SALK_0622
c	34	12.4	45.9	31	8	BH848434	BH848434 SALK_0682
	35	12.4	45.9	35	9	AJ593126	AJ593126 Arabidops
	36	12.4	45.9	52	1	AU259848	AU259848 AU259848
	37	12.4	45.9	52	7	CF353292	CF353292 lab82f08.
c	38	12.4	45.9	54	7	CF776279	CF776279 jaa17g03.
	39	12.4	45.9	54	7	CF945749	CF945749 TrEST-A38
	40	12.4	45.9	55	1	AU008564	AU008564 AU008564
c	41	12.4	45.9	56	2	AW100567	AW100567 sd57a06.y
c	42	12.4	45.9	56	8	BH866214	BH866214 SALK_1008
c	43	12.4	45.9	57	1	AA816100	AA816100 vp56e12.r
	44	12.4	45.9	58	1	AI960912	AI960912 sc92d11.y
	45	12.4	45.9	59	8	BH848238	BH848238 SALK_0676

ALIGNMENTS

RESULT 1
AU102657/c
LOCUS AU102657 50 bp mRNA linear EST 28-JAN-2004
DEFINITION AU102657 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone CAS05220, mRNA sequence.
ACCESSION AU102657
VERSION AU102657.1 GI:13552178
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CAS05220"
/clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

Query Match 53.3%; Score 14.4; DB 1; Length 50;
Best Local Similarity 66.7%; Pred. No. 4.1e+04;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 CUGACUGUUCAUCGAAACCGGGUC 26
              |:||| |   |:|||| ||||| |
Db         35 CTGACCGGGAATCGAACCCGGGCC 12
```

```
RESULT 2
BI945126/c
LOCUS       BI945126                 60 bp    mRNA    linear   EST 08-JUL-2004
DEFINITION  sb26b03.y1 Gm-c1008 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1008-294 5', mRNA sequence.
ACCESSION   BI945126
VERSION     BI945126.1  GI:16282639
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            When it has been determined, an EST from the other end of this
            clone is listed in the 'Other ESTs on clone' field. Putative full
            length read vector to vector length is 61 This clone is available
            through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA
            (phone: 800 423 4163; email: info@biogeneticservices.com)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /cultivar="Williams"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1008-294"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1008"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     cDNA library was constructed from mRNA isolated from whole
                     young pods, approximately 2cm long, of 12-week-old
                     greenhouse grown plants. The library was prepared using
                     the Life Technologies pSuperScript cDNA library
                     construction kit. Complementary DNA was synthesized from
                     mRNA using a poly (dT) sequence with a NotI restriction
                     site. SalI linkers adapters were ligated to the
                     blunt-ended cDNA fragments followed by NotI digestion. The
                     cDNA fragments were directionally cloned into the
                     NotI-SalI restriction site of the pSPORT1 vector. The
                     ligated cDNA fragments were transformed into E. coli
                     ElectroMax DH10B host cells. This library was constructed
                     by Dr. Lila Vodkin and Dr. Anu Khanna."
ORIGIN

  Query Match             53.3%;  Score 14.4;  DB 4;  Length 60;
  Best Local Similarity   54.2%;  Pred. No. 4.2e+04;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGU 25
              :|: ||:    |:|||||| |||:
Db         44 TCTCACTCAAGATCGAAACTGGGT 21


RESULT 3
AJ792078
LOCUS       AJ792078                 51 bp    mRNA    linear   EST 11-AUG-2004
DEFINITION  AJ792078 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_2_10_i05, mRNA sequence.
ACCESSION   AJ792078
VERSION     AJ792078.1  GI:51107406
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Lamiales; Plantaginaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.
  TITLE     Antirrhinum EST collection
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Antirrhinum majus"
                     /mol_type="mRNA"
                     /db_xref="taxon:4151"
                     /clone="018_2_10_i05"
                     /tissue_type="whole plant"
                     /clone_lib="Antirrhinum majus whole plant"
ORIGIN

  Query Match             51.9%;  Score 14;  DB 1;  Length 51;
  Best Local Similarity   50.0%;  Pred. No. 6.5e+04;
  Matches   11;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 UGACUGUUCAUCGAAACCGGGU 25
              :|||: ::| :|||  | |||:
Db          2 TGACTATTCGTCGACTCAGGGT 23


RESULT 4
AL765834
LOCUS       AL765834                 43 bp    DNA     linear   GSS 01-APR-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-144A04-012978,
            genomic survey sequence.
ACCESSION   AL765834
VERSION     AL765834.1  GI:21518967
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
  MEDLINE   22755829
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
  MEDLINE   23117147
   PUBMED   14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
```

```
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
   PUBMED   14682050
REFERENCE   4  (bases 1 to 43)
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion within the locus defined by BAC clone
            f25p17. Details on the protocols used for generation of the
            sequence are described in References 1-3. The sequences are
            generated at the MPI for Plant Breeding Research in the context of
            the GABI-Kat project. GABI-Kat is part of the German Plant Genomics
            program designated 'GABI'. Information on line availability can be
            found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-144A04-012978"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             51.1%;  Score 13.8;  DB 9;  Length 43;
  Best Local Similarity   44.0%;  Pred. No. 7.9e+04;
  Matches   11;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUC 26
              :|:| |:|::  :  |||  |||:|
Db          6 TCTGTCTGTTGTTTAAAAGTGGGTC 30


RESULT 5
AV834227/c
LOCUS       AV834227                 48 bp    mRNA    linear   EST 09-MAY-2002
DEFINITION  AV834227 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
            vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA
            clone bags7j23, mRNA sequence.
ACCESSION   AV834227
VERSION     AV834227.1  GI:14526316
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Sato,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kazsato@rib.okayama-u.ac.jp,
            URL:http://www.rib.okayama-u.ac.jp/barley/
            database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="Haruna Nijo"
                     /sub_species="vulgare"
                     /db_xref="taxon:112509"
                     /clone="bags7j23"
                     /tissue_type="shoots"
                     /dev_stage="germination"
                     /clone_lib="K. Sato unpublished cDNA library: Hordeum
                     vulgare subsp. vulgare shoots germination"
ORIGIN

  Query Match             51.1%;  Score 13.8;  DB 1;  Length 48;
  Best Local Similarity   53.8%;  Pred. No. 8e+04;
  Matches   14;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 GUCUGACUGUUCAUCGAAACCGGGUC 26
              |:| | | |: ||: |  |||| |:|
Db         34 GTCCGNCCGTCCATGGCTACCGTGTC 9


RESULT 6
BH896046
LOCUS       BH896046                 49 bp    DNA     linear   GSS 14-AUG-2002
DEFINITION  3526_1_36_1_E08.2EL_y_1 3526 - RescueMu Grid K Zea mays genomic,
            genomic survey sequence.
ACCESSION   BH896046
VERSION     BH896046.1  GI:22230589
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 3526_1_36_1  row: 11
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="3526 - RescueMu Grid K"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid K was grown at Molokai, Hawaii in Winter
                     2000-2001. DNA was extracted from leaf punches, double
                     digested using BamHI and BglII, and ligated to form
                     circular plasmids. DH10B cells were transformed and then
                     screened on LB plates with ampicillin."
ORIGIN

  Query Match             51.1%;  Score 13.8;  DB 8;  Length 49;
  Best Local Similarity   48.0%;  Pred. No. 8e+04;
  Matches   12;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;
```

```
Qy          3 CUGACUGUUCAUCGAAACCGGGUCC 27
              |:| |: ::||:| |   | ||:||
Db         10 CTGCCTCTTCATCCACCTCTGGTCC 34


RESULT 7
AJ593160/c
LOCUS       AJ593160                 52 bp    DNA     linear   GSS 15-JAN-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            376A08, genomic survey sequence.
ACCESSION   AJ593160
VERSION     AJ593160.1  GI:37942784
KEYWORDS    GSS; left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 52)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="376A08"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .52
                     /note="T-DNA flanking sequence
                     left border"
ORIGIN

  Query Match           51.1%;  Score 13.8;  DB 9;  Length 52;
  Best Local Similarity   52.9%;  Pred. No. 8.1e+04;
  Matches    9;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAA 18
              :|: ||:|::||:| ||
Db         40 TCTCACTGTTCATCCAA 24


RESULT 8
AZ475845
LOCUS       AZ475845                 23 bp    DNA     linear   GSS 04-OCT-2000
DEFINITION  1M0294P09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0294P09 F, genomic survey sequence.
ACCESSION   AZ475845
VERSION     AZ475845.1  GI:10633970
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0294  row: P  column: 09
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 23.
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0294P09"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match           50.4%;  Score 13.6;  DB 8;  Length 23;
  Best Local Similarity   60.0%;  Pred. No. 9.1e+04;
  Matches   12;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 CUGACUGUUCAUCGAAACCG 22
              |:|||:|::||  || || |
Db          2 CTGACTGTTCACAGACACAG 21


RESULT 9
CL314996/c
LOCUS       CL314996                 60 bp    DNA     linear   GSS 01-MAR-2004
DEFINITION  mth2-133P14_OP Medicago truncatula BAC end sequences Medicago
            truncatula genomic 5', genomic survey sequence.
ACCESSION   CL314996
VERSION     CL314996.1  GI:44832670
KEYWORDS    GSS.
```

```
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Jakab,J., Deak,G., Kevei,Z., Karchesz,K., Sarai,E., Kiss,P.,
            Kereszt,A., Kalo,P., Endre,G. and Kiss,G.B.
  TITLE     Medicago Truncatula BAC end sequencing
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Deak,G.
            Alfalfa Genomics Group; Medicago Genetics Group
            Agricultural Biotechnology Center; Biological Research Center
            P.O. Box 411, Hungary, H-2100 Godollo, Szent-Gyorgyi Albert ut 4.;
            P.O. BOX 521, Hungary, H-6701 Szeged, Temesvari krt. 62
            Tel: 3628526142
            Fax: 3628526193
            Email: gdeak@abc.hu
            Plate: 133  row: P  column: 14
            Seq primer: OP Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /cultivar="Jemalong"
                     /isolate="A17"
                     /db_xref="taxon:3880"
                     /sex="Hermaphrodite"
                     /clone_lib="Medicago truncatula BAC end sequences"
                     /note="Organ: Leaf; Vector: pBelo11; Site_1: HindIII;
                     Site_2: HindIII; Construction of a bacterial artificial
                     chromosome library of Medicago truncatula and
                     identification of clones containing ethylene-response
                     genes. Theor Appl Genet (1999) 98: 638-646 Y.-W.,Nam;
                     R.V.,Penmetsa; G.,Endre; P.,Uribe; D.,Kim; D.R.,Cook"
ORIGIN

  Query Match             49.6%;  Score 13.4;  DB 9;  Length 60;
  Best Local Similarity   52.2%;  Pred. No. 1.3e+05;
  Matches   12;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GUCUGACUGUUCAUCGAAACCGG 23
              |:|:| |  ::||:  |||| ||
Db         47 GTCTGCCCATTCATACAAACAGG 25


RESULT 10
CL265635
LOCUS       CL265635                  43 bp    DNA     linear   GSS 03-FEB-2004
DEFINITION  02S0288-04B1-A02 UniformMu MuTAIL Library Zea mays genomic clone
            02S0288-04B1-A02, genomic survey sequence.
ACCESSION   CL265635
VERSION     CL265635.1  GI:41873394
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Latshaw,S., Tan,B-.C., Settles,A.M. and McCarty,D.R.
  TITLE     Sequence tagged transposon insertions from the UniformMu maize
            population
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Donald R. McCarty
            Plant Molecular and Cellular Biology Program
            University of Florida
            PO 110690 Gainesville, FL 32611-0690, USA
            Tel: 352-392-1928 x322
            Email: drm@ufl.edu
            Sequence flanking probable Mu insertion site in UniformMu line:
            02S0288-04, Primer set: B
            Class: transposon insertion site.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="W22 (ACR, bz1-m9)"
                     /cultivar="UniformMu"
                     /db_xref="taxon:4577"
                     /clone="02S0288-04B1-A02"
                     /clone_lib="UniformMu MuTAIL Library"
                     /note="Vector: TOPO-PCR4; DNA flanking Mu transposon
                     insertions in Mu inactive lines were extracted from the
                     UniformMu maize population by the thermo assymmetric
                     interlaced PCR (TAIL) protocol using primers specific for
                     the Mu terminal inverted repeat and a set of 16 arbitrary
                     primers. Amplicons were size enriched using Sepharose 400
                     spin columns and cloned into the TOPO PCR4 vector."
ORIGIN

  Query Match             48.9%;  Score 13.2;  DB 9;  Length 43;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+05;
  Matches   13;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 GUCUGACUGUUCAUCGAAACCGGGUC 26
              | |:  |  ::| :|||| |||| :|
Db          3 GCCTTCCATTTCGTCGAATCCGGTTC 28


RESULT 11
AZ820793/c
LOCUS       AZ820793                  45 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0093B18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0093B18 F, genomic survey sequence.
ACCESSION   AZ820793
VERSION     AZ820793.1  GI:12990701
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0093  row: B  column: 18
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 45.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0093B18"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
```

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

```
  Query Match             48.9%;  Score 13.2;  DB 8;  Length 45;
  Best Local Similarity   53.8%;  Pred. No. 1.6e+05;
  Matches   14;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 GUCUGACUGUUCAUCGAAACCGGGUC 26
              |:|:  |:|  ||  ||||||  |:|
Db         41 GTCTATCTGCACAGTGAAAACCCTGTC 16
```

RESULT 12
AI119755
LOCUS AI119755 46 bp mRNA linear EST 02-SEP-1998
DEFINITION uc20a09.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone IMAGE:1398520 5' similar to WP:M02B7.4 CE12326 ;, mRNA sequence.
ACCESSION AI119755
VERSION AI119755.1 GI:3520079
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 46)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:910236
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .46
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1398520"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares_mammary_gland_NbMMG"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

ORIGIN

```
  Query Match             48.9%;  Score 13.2;  DB 1;  Length 46;
  Best Local Similarity   46.2%;  Pred. No. 1.6e+05;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUCC 27
              :| |||: :: |:||  | || :||
Db          9 TCAGACTCTTGATCGTGGCTGGATCC 34
```

RESULT 13
TA126C12Q
LOCUS TA126C12Q 53 bp DNA linear GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 126c12, reverse sequence, genomic survey sequence.
ACCESSION AL463403
VERSION AL463403.1 GI:11833913
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 53)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh1@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
source 1. .53
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="126c12"

ORIGIN

```
  Query Match             48.9%;  Score 13.2;  DB 9;  Length 53;
  Best Local Similarity   46.2%;  Pred. No. 1.6e+05;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUCC 27
              :|: | : : ||:  ||  ||||:||
Db         14 TCTTAATATACATTAAATTCGGGTCC 39
```

RESULT 14
AA905188/c
LOCUS AA905188 43 bp mRNA linear EST 09-JUN-1998
DEFINITION ok10b03.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1507373 3' similar to SW:PDXK_HUMAN O00764 PYRIDOXINE KINASE ;, mRNA sequence.
ACCESSION AA905188
VERSION AA905188.1 GI:3040311
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 43)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 530 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .43
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1507373"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

Query Match 48.1%; Score 13; DB 1; Length 43;
Best Local Similarity 47.6%; Pred. No. 1.9e+05;
Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 GUCUGACUGUUCAUCGAAACC 21
              |:|:  |:|:: |: ||| ||
Db         34 GTCTCCCTGTTGATGGAACCC 14
```

RESULT 15
BZ594864
LOCUS BZ594864 44 bp DNA linear GSS 07-JAN-2003
DEFINITION SALK_085323.43.20.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_085323.43.20.x, genomic survey sequence.
ACCESSION BZ594864
VERSION BZ594864.1 GI:27535383
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 44)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At5g35690.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .44
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_085323.43.20.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
ORIGIN

Query Match 48.1%; Score 13; DB 8; Length 44;
Best Local Similarity 52.4%; Pred. No. 1.9e+05;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

```
Qy          5 GACUGUUCAUCGAAACCGGGU 25
              || :|::||:|  ||  |||:
Db         13 GAGTGTTCATCAGAAAAGGGT 33
```

Search completed: October 31, 2004, 04:14:10
Job time : 1373.36 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 04:14:22 ; Search time 1728 Seconds
(without alignments)
1855.727 Million cell updates/sec

Title: US-09-974-974-5
Perfect score: 88
Sequence: 1 accguugguuuccguagugu.....cgggcacuacaaaaaccaac 88

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 3813152

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40.4	45.9	57	2	BF031637	BF031637 601558122
c 2	39	44.3	158	8	CC059191	CC059191 ii20e02.b
3	38.2	43.4	171	9	AJ591533	AJ591533 Arabidops
4	33	37.5	174	8	CC004711	CC004711 PUDGC26TD
5	31.6	35.9	177	9	BX176513	BX176513 Danio rer
6	30.8	35.0	126	7	H04713	H04713 e17-1.3.3-3
7	29.8	33.9	71	9	BX203515	BX203515 Danio rer
8	29.6	33.6	150	5	BU275850	BU275850 Cr_Emb_08
9	29.4	33.4	117	6	CD768454	CD768454 AGENCOURT
c 10	29.2	33.2	130	4	BI845249	BI845249 fs92a07.y
11	29.2	33.2	166	8	BH330762	BH330762 CH230-14I
c 12	29	33.0	157	8	AZ261468	AZ261468 RPCI-23-1
13	29	33.0	166	8	AZ118437	AZ118437 RPCI-23-4
14	29	33.0	175	9	BX244513	BX244513 Danio rer
15	28.8	32.7	158	7	CO739878	CO739878 SlLB06a25
16	28.8	32.7	166	7	CO739859	CO739859 SlLB06a25
c 17	28.4	32.3	158	4	BI708958	BI708958 fp94g05.y
18	28.4	32.3	169	8	BH812737	BH812737 SALK_0629
c 19	28.4	32.3	195	4	BM532607	BM532607 fx38c08.y
20	28.2	32.0	108	7	CO739537	CO739537 SlLB06a24
21	28.2	32.0	118	7	CO739528	CO739528 SlLB06a23
22	28.2	32.0	129	7	CO739710	CO739710 SlLB06a24
23	28.2	32.0	129	7	CO739745	CO739745 SlLB06a24
24	28.2	32.0	132	7	CO739585	CO739585 SlLB06a24
25	28.2	32.0	133	7	CO739781	CO739781 SlLB06a24
26	28.2	32.0	134	7	CO739672	CO739672 SlLB06a24
27	28.2	32.0	135	7	CO739592	CO739592 SlLB06a24
28	28.2	32.0	135	7	CO739661	CO739661 SlLB06a24
29	28.2	32.0	135	7	CO739756	CO739756 SlLB06a24
30	28.2	32.0	135	7	CO740043	CO740043 SlLB06a25
31	28.2	32.0	136	7	CO739581	CO739581 SlLB06a24
32	28.2	32.0	136	7	CO739594	CO739594 SlLB06a24
33	28.2	32.0	136	7	CO739620	CO739620 SlLB06a24
34	28.2	32.0	136	7	CO739681	CO739681 SlLB06a24
35	28.2	32.0	136	7	CO739794	CO739794 SlLB06a24
36	28.2	32.0	136	7	CO739992	CO739992 SlLB06a25
37	28.2	32.0	137	7	CO739512	CO739512 SlLB06a23
38	28.2	32.0	137	7	CO739733	CO739733 SlLB06a24
39	28.2	32.0	137	7	CO739810	CO739810 SlLB06a25
40	28.2	32.0	137	7	CO740001	CO740001 SlLB06a25
41	28.2	32.0	138	7	CO739739	CO739739 SlLB06a24
42	28.2	32.0	138	7	CO739925	CO739925 SlLB06a25
43	28.2	32.0	139	7	CO739747	CO739747 SlLB06a24
44	28.2	32.0	139	7	CO739827	CO739827 SlLB06a25
45	28.2	32.0	140	7	CO739573	CO739573 SlLB06a24

ALIGNMENTS

RESULT 1
BF031637
LOCUS BF031637 57 bp mRNA linear EST 10-OCT-2000
DEFINITION 601558122F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3827860 5', mRNA sequence.
ACCESSION BF031637
VERSION BF031637.1 GI:10739349
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 57)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM497 row: 1 column: 05
High quality sequence stop: 57.
FEATURES Location/Qualifiers
source 1. .57
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3827860"
/tissue_type="hypernephroma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_58"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```
ORIGIN

  Query Match             45.9%;  Score 40.4;  DB 2;  Length 57;
  Best Local Similarity   66.0%;  Pred. No. 0.00023;
  Matches   33;  Conservative   11;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         18 UGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAA 67
              :|:|| ||::|:||| ::|||| |||||||||||||:||||||::   ||
Db          1 TGTAGCGGTTATCACATTCGCCTCACACGCGAAAGGTCCCCGGTTTATAA 50


RESULT 2
CC059191/c
LOCUS       CC059191                158 bp    DNA     linear   GSS 03-APR-2003
DEFINITION  ii20e02.b1 WGS-SbicolorF (DH5a methyl filtered) Sorghum bicolor
            genomic clone ii20e02, genomic survey sequence.
ACCESSION   CC059191
VERSION     CC059191.1  GI:29503921
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 158)
  AUTHORS   Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
            Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
            Zutavern,T., Palmer,L., McCombie,W.R. and Martienssen,R.A.
  TITLE     Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
            Plate: ii20  row: e  column: 02
            Seq primer: -21M13UnivFwd
            Class: shotgun
            High quality sequence stop: 158.
FEATURES             Location/Qualifiers
     source          1. .158
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4558"
                     /clone="ii20e02"
                     /lab_host="DH5a"
                     /clone_lib="WGS-SbicolorF (DH5a methyl filtered)"
                     /note="Site_1: Xba I; Site_2: Xba I; The vector was
                     digested with XbaI and one nucleotide was added by fill in
                     in the recessive 3' end. The genomic DNA was nebulized,
                     end repaired, adaptor ligated and size fractionated using
                     sephadex. The resulting fragments were between 0.8 and 3
                     kb and were cloned into the vector (.x/y reads in M13mp19,
                     .b/g reads in pUC19). The same ligation was transformed
                     into DH5a."

ORIGIN

  Query Match             44.3%;  Score 39;  DB 8;  Length 158;
  Best Local Similarity   61.9%;  Pred. No. 0.0009;
  Matches   39;  Conservative    9;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         23 UGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGCACUACAAAA 82
              :||::|:||||:  | |:|||||| |  ||||:| ||||::|||| ||||||      |||
Db         78 TGGTTATCACGTCAGTCTAACACACTGAAGGTCTCCGGTTCGAACCCGGGCGACGGCAAA 19

Qy         83 ACC 85
              | |
Db         18 ATC 16


RESULT 3
```

```
AJ591533
LOCUS       AJ591533                171 bp    DNA     linear   GSS 15-JAN-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            587A06, genomic survey sequence.
ACCESSION   AJ591533
VERSION     AJ591533.1  GI:37941157
KEYWORDS    GSS; left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 171)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .171
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="587A06"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .171
                     /note="T-DNA flanking sequence
                     left border"

ORIGIN

  Query Match             43.4%;  Score 38.2;  DB 9;  Length 171;
  Best Local Similarity   55.2%;  Pred. No. 0.0018;
  Matches   37;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          7 GGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAA 66
              ||:::|          :||::|:||||:  | |:||||| |  ||||:| ||||::||||
Db         16 GGTTTCGTGGTGTAGTTGGTTATCACGTCAGTCTAACACACTGAAGGTCTCCGGTTCGAA 75

Qy         67 ACCGGGC 73
              |||||||
Db         76 CCCGGGC 82


RESULT 4
CC004711
LOCUS       CC004711                174 bp    DNA     linear   GSS 31-MAR-2003
DEFINITION  PUDGC26TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa179F03,
            genomic survey sequence.
ACCESSION   CC004711
VERSION     CC004711.1  GI:29383271
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 174)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .174
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa179F03"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             37.5%;  Score 33;  DB 8;  Length 174;
  Best Local Similarity   57.5%;  Pred. No. 0.17;
  Matches   42;  Conservative    6;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         14 GUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGC 73
              |:||:  ||  | :|  ||| :||| :| || |||| |||  |  || :||| |||  ||
Db         66 GTAGTTCAGATGGTAGAACGCTCGCTTAGCAAGCGAGAGGCGCGGGGATCGATACCCCGC 125

Qy         74 ACUACAAAAACCA 86
              |  | |||| ||
Db        126 ATCTCCAAAATCA 138


RESULT 5
BX176513
LOCUS       BX176513                177 bp    DNA     linear   GSS 13-MAR-2003
DEFINITION  Danio rerio genomic clone DKEY-176E1, genomic survey sequence.
ACCESSION   BX176513
VERSION     BX176513.1  GI:28008239
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 177)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the T7 end of BAC 176E1. 176E1 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .177
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-176E1"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN

  Query Match             35.9%;  Score 31.6;  DB 9;  Length 177;
  Best Local Similarity   54.8%;  Pred. No. 0.57;
  Matches   34;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         12 CCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGG 71
              | |:| :| ||:||  | ||||::||||: |||     ||||:| | ||:: ||  |  |
Db        115 CAGTAATGCAGTGGGCAGCACGTTCGCCTCACAGCAAGAAGGTCGCTGGTTTGAGCCTTG 174

Qy         72 GC 73
              ||
Db        175 GC 176


RESULT 6
H04713
LOCUS       H04713                  126 bp    mRNA    linear   EST 20-JUN-1995
DEFINITION  e17-1.3.3-3' Lambda Zap Express library of P. E. Schwartz (9/93)
            Rattus norvegicus cDNA clone e17-1.3.3 3', mRNA sequence.
ACCESSION   H04713
VERSION     H04713.1  GI:867646
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 126)
  AUTHORS   Schwartz,P.E., Grieshaber,N.A., Grieshaber,S.S. and Majack,R.A.
  TITLE     An expressed sequence tag from in vitro embryonic rat vascular
            smooth muscle cells
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Phillip E. Schwartz
            Pediatrics and Cellular and Structural Biology
            University of Colorado Health Sciences Center
            4200 East Ninth Avenue, Denver, Colorado, 80262, USA
            Tel: (303) 270-4569
            Fax: (303) 270-8353
            Email: schwartz_p@defiance.hsc.colorado.edu.
FEATURES             Location/Qualifiers
     source          1. .126
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="e17-1.3.3"
                     /clone_lib="Lambda Zap Express library of P. E. Schwartz
                     (9/93)"
                     /note="developmental-stage=Embryonic Day 17
                     Post-fertilization; tissue-type=Aorta; cell-type=Vascular
                     smooth muscle cell; sex=Male."
ORIGIN

  Query Match             35.0%;  Score 30.8;  DB 7;  Length 126;
  Best Local Similarity   52.7%;  Pred. No. 1.1;
  Matches   39;  Conservative    8;  Mismatches   27;  Indels    0;  Gaps    0;

Qy         15 UAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGCA 74
              :||   ||:||:     || : |||:|  | ||| |||| ||  ||::||   ||  ||
Db         51 TAGCTCAGTGGTAGAGGCGCTTGCCTAGGAAGCGCAAGGCCCTGGGTTCGGTCCCCAGCT 110

Qy         75 CUACAAAAACCAAC 88
              |:  |||||  |||
Db        111 CTGAAAAAAAGAAC 124


RESULT 7
BX203515
LOCUS       BX203515                71 bp    DNA     linear   GSS 13-MAR-2003
DEFINITION  Danio rerio genomic clone DKEY-224D6, genomic survey sequence.
ACCESSION   BX203515
VERSION     BX203515.1  GI:28035401
KEYWORDS    GSS.
```

```
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the T7 end of BAC 224D6. 224D6 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-224D6"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN

  Query Match             33.9%;  Score 29.8;  DB 9;  Length 71;
  Best Local Similarity   49.2%;  Pred. No. 2.4;
  Matches   32;  Conservative   11;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          9 UUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAAC 68
              ::    |: |    ||:||::|   ||  :||||: |||     |||||:| | ||::  || |
Db          2 TTGGAGTGGCTCAGTGGTTAGTACCGTCGCCTCACAGTAAAAAGGTCACTGGTTTAAATC 61

Qy         69 CGGGC 73
                | |||
Db         62 CCGGC 66


RESULT 8
BU275850
LOCUS       BU275850                150 bp    mRNA    linear   EST 03-JAN-2003
DEFINITION  Cr_Emb_08F08_TEXF1 Convoluta roscoffensis embryos from Eva Jiminez
            Symsagittifera roscofensis cDNA clone Cr_Emb_08F08 5', mRNA
            sequence.
ACCESSION   BU275850
VERSION     BU275850.1  GI:27476620
KEYWORDS    EST.
SOURCE      Symsagittifera roscofensis
  ORGANISM  Symsagittifera roscofensis
            Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Acoelomorpha;
            Acoela; Sagittiferidae; Symsagittifera.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Blaxter,M.L., Jiminez,E., Rosie,A., Aboobaker,A. and Salo,E.
  TITLE     A survey of genes expressed in the acoel flatworm Convoluta
            roscoffensis
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Blaxter ML
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
            3JT, UK.
            Tel: +44 131 650 6760
            Fax: +44 131 670 5450
            Email: mark.blaxter@ed.ac.uk
            The library was prepared by Eva Jimenez, University of Barcelona,
            Spain The sequencing was performed by Eva Jimenez, Aziz Aboobaker
            and Ailie Rosie Edinburgh Phylogenomics Programme, ICAPB,
            Edinburgh, UK.The sequence contained a PolyA tail (trimmed)
            PCR PRimers
            FORWARD: T7PL
            BACKWARD: triple 5 Exseq
            Plate: 08  row: F  column: 08
            Seq primer: TEXF1
            High quality sequence stop: 150.
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="Symsagittifera roscofensis"
                     /mol_type="mRNA"
                     /db_xref="taxon:84072"
                     /clone="Cr_Emb_08F08"
                     /sex="mixed"
                     /dev_stage="embryo"
                     /clone_lib="Convoluta roscoffensis embryos from Eva
                     Jiminez"
                     /note="Vector: lambda triplex; Site_1: EcoRI; Site_2:
                     HinDIII; The library was prepared by Eva Jimenez,
                     University of Barcelona, Spain The sequencing was
                     performed by Eva Jimenez,  Aziz Aboobaker and Ailie Rosie,
                     Edinburgh Phylogenomics Programme,  ICAPB, Edinburgh, UK."
ORIGIN

  Query Match             33.6%;  Score 29.6;  DB 5;  Length 150;
  Best Local Similarity   50.0%;  Pred. No. 3.2;
  Matches   34;  Conservative   10;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          7 GGUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAA 66
              ||    | |: |:| ||:||::| |    : ||||:  ||  |||||||:|     |::|||
Db         26 GGGAGCAGTGGTGAAGTGGTTAACGTCTCCGCCTCTCAATCGAAAGGTCGGATGTTCGAG 85

Qy         67 ACCGGGCA 74
                ||   ||
Db         86 GCCCACCA 93


RESULT 9
CD768454
LOCUS       CD768454                117 bp    mRNA    linear   EST 02-JUL-2003
DEFINITION  AGENCOURT_14739547 NICHD_MM_Hyp1 Mus musculus cDNA clone
            IMAGE:6975678 5', mRNA sequence.
ACCESSION   CD768454
VERSION     CD768454.1  GI:32426956
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 117)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Pamela Mellon and Dr. Pat Chappell
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3276  row: k  column: 05
            High quality sequence stop: 117.
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6975678"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_MM_Hyp1"
                     /note="Organ: brain - normal mediobasal hypothalamus;
                     Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc); Site_2:
                     SfiI (ggccgcctcggcc); Non-normalized full-length enriched
                     library 5' and 3' adaptors were used in cloning as
```

follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.08 kb (range .73-1.37 kb). 13/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA).Corp."

ORIGIN

Query Match 33.4%; Score 29.4; DB 6; Length 117;
Best Local Similarity 57.1%; Pred. No. 3.6;
Matches 36; Conservative 6; Mismatches 21; Indels 0; Gaps 0;

```
Qy         25 GUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAAC 84
              | :|    || : |||:| || ||| |||| ||  ||::||   |    || |:|||||||
Db         28 GGTAGAGCGCTTGCCTAGCAAGCGCAAGGCCCTGGGTTCGGTCCTCAGCTCTACAAAAAA 87

Qy         85 CAA 87
              ||
Db         88 AAA 90
```

RESULT 10
BI845249/c
LOCUS BI845249 130 bp mRNA linear EST 04-OCT-2001
DEFINITION fs92a07.y1 Zebrafish neuronal Danio rerio cDNA clone IMAGE:5077140 5', mRNA sequence.
ACCESSION BI845249
VERSION BI845249.1 GI:15957772
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 130)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library constructed by S. Lin DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: T3 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .130
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5077140"
/sex="mixed"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Zebrafish neuronal"
/note="Organ: brain; Vector: pBluescript (modified); Site_1: DraIII(Y); Site_2: DRAIII(X); Library is cloned directionally between the DraIII(X) and DraIII(Y) sites and has been amplified. Library constructed by S. Lin. "
ORIGIN

Query Match 33.2%; Score 29.2; DB 4; Length 130;
Best Local Similarity 53.4%; Pred. No. 4.4;
Matches 31; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

```
Qy          8 GUUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGA 65
              |:  | |: | | ||:|| :| |||| :||||: |||     ||| :| | ||::|||
Db         63 GTCACGGTGGCGCAGTGGGTAGCACGATCGCCTCACAGCAAGAAGATCACTGGTTCGA 6
```

RESULT 11
BH330762
LOCUS BH330762 166 bp DNA linear GSS 03-DEC-2001
DEFINITION CH230-14I7.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone CH230-14I7, genomic survey sequence.
ACCESSION BH330762
VERSION BH330762.1 GI:17261476
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 166)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: CH230-14I7.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or ering_information.htm). BAC end page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 14 row: I column: 7
Seq primer: SP6
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .166
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-14I7"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI; CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by Pieter de Jong"
ORIGIN

Query Match 33.2%; Score 29.2; DB 8; Length 166;
Best Local Similarity 52.7%; Pred. No. 4.5;
Matches 39; Conservative 7; Mismatches 28; Indels 0; Gaps 0;

```
Qy         15 UAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGCA 74
              :||   ||:||:     || :  ||:|  | ||| |||| ||  ||::||| ||  ||
Db         69 TAGCTCAGTGGTAGAAGCGCTTACCTAGGAAGCGCAAGGCCCTGGGTTCGATCCCCAGCT 128

Qy         75 CUACAAAAACCAAC 88
              |  |||||  |||
Db        129 CCGAAAAAAAGAAC 142
```

RESULT 12
AZ261468/c

```
LOCUS       AZ261468                157 bp    DNA     linear   GSS 26-JUL-2000
DEFINITION  RPCI-23-153F14.TJ RPCI-23 Mus musculus genomic clone
            RPCI-23-153F14, genomic survey sequence.
ACCESSION   AZ261468
VERSION     AZ261468.1  GI:9469876
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 157)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-153F14.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 153  row: F  column: 14
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .157
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-153F14"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match           33.0%;  Score 29;  DB 8;  Length 157;
  Best Local Similarity  54.1%;  Pred. No. 5.4;
  Matches   33;  Conservative    8;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         26 UUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACC 85
              ::|   |   :: |||:| || ||  || | ||  ||::| |    |  ||||:||| |||||
Db        132 TTAAAATATTTGCCTAGCATGCACAAAGCCCTGGGTTCAATCTCTAGCACTACATAAACC 73

Qy         86 A 86
              |
Db         72 A 72


RESULT 13
AZ118437
LOCUS       AZ118437                166 bp    DNA     linear   GSS 12-MAY-2000
DEFINITION  RPCI-23-476F9.TJ RPCI-23 Mus musculus genomic clone RPCI-23-476F9,
            genomic survey sequence.
ACCESSION   AZ118437
VERSION     AZ118437.1  GI:7782376
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 166)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-476F9.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 476  row: F  column: 9
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .166
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-476F9"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match           33.0%;  Score 29;  DB 8;  Length 166;
  Best Local Similarity  52.2%;  Pred. No. 5.4;
  Matches   36;  Conservative    8;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         15 UAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGCA 74
              :||:  ||: | :| || | :| ||:|||| ||  |  | ||  ||:: ||  |  | ||
Db         89 TAGTTCAGTTGATAGCATGCTCTCCTAACATGCACATAGCCCTGGGTTTGATCCTCAGAA 148

Qy         75 CUACAAAAA 83
              | ||| |||
Db        149 CAACATAAA 157


RESULT 14
BX244513
LOCUS       BX244513                175 bp    DNA     linear   GSS 13-MAR-2003
DEFINITION  Danio rerio genomic clone DKEY-233N3, genomic survey sequence.
ACCESSION   BX244513
VERSION     BX244513.1  GI:28166847
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 175)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the T7 end of BAC 233N3. 233N3 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .175
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-233N3"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN

  Query Match             33.0%;  Score 29;  DB 9;  Length 175;
  Best Local Similarity   50.7%;  Pred. No. 5.4;
  Matches   35;  Conservative    9;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         12 CCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGG 71
              || : |   ||:||::| |||  :||| : |||      ||||:||| ||::  |  || |
Db         42 CCATGGCTCAGTGGTTAGCACTGTCGCATCACAACAAGAAGGTCCCTGGTTTAAGTCCCG 101

Qy         72 GCACUACAA 80
              ||   || |
Db        102 GCTGGACGA 110


RESULT 15
CO739878
LOCUS       CO739878                 158 bp    mRNA    linear   EST 29-JUL-2004
DEFINITION  SlLB06a25e03f1 squirrel brain library 1 Spermophilus lateralis cDNA
            clone 25e03 5', mRNA sequence.
ACCESSION   CO739878
VERSION     CO739878.1  GI:50827148
KEYWORDS    EST.
SOURCE      Spermophilus lateralis (golden-mantled ground squirrel)
  ORGANISM  Spermophilus lateralis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
            Spermophilus.
REFERENCE   1  (bases 1 to 158)
  AUTHORS   Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
            Rogers,J. and Cossins,A.R.
  TITLE     Microarray analysis of transcriptional changes during hibernation
            in the golden mantled ground squirrel, Spermophilus lateralis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Andrew R. Cossins
            Laboratory for Environmental Gene Regulation
            University of Liverpool
            School of Biological Sciences, The Biosciences Building, Crown
            Street, Liverpool, United Kingdom, L69 7ZB
            Tel: +44(0)151-795-4510
            Fax: +44(0)151-795-4431
            Email: cossins@liv.ac.uk
            Vector has been trimmed from this EST.
            Plate: 25  row: e  column: 03
            Seq primer: TriplEx 5'LD (5'-CTCGGGAAGCGCGCCATTGTGTTGGT-3')
            High quality sequence stop: 158.
FEATURES             Location/Qualifiers
     source          1. .158
                     /organism="Spermophilus lateralis"
                     /mol_type="mRNA"
                     /db_xref="taxon:76772"
                     /clone="25e03"
                     /sex="Male & female"
                     /tissue_type="brain"
                     /dev_stage="Adult"
                     /lab_host="E.coli Electromax DH10B"
                     /clone_lib="squirrel brain library 1"
                     /note="Vector: pTriplEx2; Site_1: SfiI GGCCATTACGGCC;
                     Site_2: SfiI GGCCGCCTCGGCC; Normalized cDNA library
                     prepared from brain of hibernating and summer animals"
ORIGIN

  Query Match             32.7%;  Score 28.8;  DB 7;  Length 158;
  Best Local Similarity   56.2%;  Pred. No. 6.4;
  Matches   36;  Conservative    6;  Mismatches   22;  Indels    0;  Gaps    0;

Qy         25 GUUAUCACGUUCGCCUAACACGCGAAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAAC 84
              | :|   || : |||:| || ||| |||| ||  ||::||   |   || |:  |||||
Db         16 GGTAGAGCGCTTGCCTAGCATGCGTAAGGCCCTGGGTTCGGTCCTCAGCCCTGTAAAAAA 75

Qy         85 CAAC 88
              ||||
Db         76 CAAC 79


Search completed: October 31, 2004, 05:19:40
Job time : 1731 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 01:51:44 ; Search time 852.61 Seconds
(without alignments)
1774.868 Million cell updates/sec

Title: US-09-974-974-1
Perfect score: 32
Sequence: 1 gguccuggccugaugagagugaugagcucuuc 32

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 1924030

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	32	100.0	32	6	AX453842	AX453842 Sequence
	2	32	100.0	32	6	BD143498	BD143498 Nucleic a
	3	21.4	66.9	33	6	AX453844	AX453844 Sequence
	4	21.4	66.9	33	6	BD143500	BD143500 Nucleic a
	5	17.8	55.6	60	6	CQ547198	CQ547198 Sequence
	6	17.4	54.4	29	6	BD197186	BD197186 Method an
	7	17.4	54.4	40	6	AX453853	AX453853 Sequence
c	8	17.4	54.4	47	6	AR291770	AR291770 Sequence
	9	16.6	51.9	30	6	AR179707	AR179707 Sequence
	10	16.6	51.9	30	6	BD222050	BD222050 SH2 domai
c	11	16.4	51.2	27	6	BD162072	BD162072 Method fo
	12	16.4	51.2	55	6	AX211297	AX211297 Sequence
	13	16.2	50.6	39	6	AR213624	AR213624 Sequence
	14	16.2	50.6	39	6	BD057682	BD057682 Fusion pr
	15	16.2	50.6	39	6	BD081512	BD081512 Soluble s
c	16	16.2	50.6	51	6	AR145871	AR145871 Sequence
c	17	16.2	50.6	60	6	CQ550477	CQ550477 Sequence
	18	16	50.0	29	6	BD199490	BD199490 Method an
	19	16	50.0	53	6	AR105779	AR105779 Sequence
c	20	15.8	49.4	38	6	AX218725	AX218725 Sequence
	21	15.6	48.8	50	6	AX161018	AX161018 Sequence
	22	15.6	48.8	60	6	CQ553925	CQ553925 Sequence
c	23	15.4	48.1	25	6	AX533469	AX533469 Sequence
c	24	15.4	48.1	25	6	AX533472	AX533472 Sequence
	25	15.4	48.1	36	6	AR041983	AR041983 Sequence
	26	15.4	48.1	36	6	AX637441	AX637441 Sequence
	27	15.4	48.1	38	6	AR334837	AR334837 Sequence
c	28	15.4	48.1	51	6	AX190288	AX190288 Sequence
	29	15.4	48.1	57	4	BOVOST	M31318 Bovine oste
	30	15.4	48.1	60	6	CQ535421	CQ535421 Sequence
c	31	15.2	47.5	28	6	I33133	I33133 Sequence 3
c	32	15.2	47.5	60	6	CQ551175	CQ551175 Sequence
c	33	15	46.9	25	6	AX533467	AX533467 Sequence
c	34	15	46.9	25	6	AX533468	AX533468 Sequence
	35	15	46.9	27	6	BD206476	BD206476 Enzymatic
	36	15	46.9	29	6	BD202230	BD202230 Method an
c	37	15	46.9	50	6	CQ008519	CQ008519 Sequence
	38	15	46.9	51	6	CQ007060	CQ007060 Sequence
c	39	15	46.9	58	6	AX924083	AX924083 Sequence
c	40	15	46.9	60	6	CQ538656	CQ538656 Sequence
c	41	14.8	46.2	34	6	I29828	I29828 Sequence 14
c	42	14.8	46.2	38	6	AR331695	AR331695 Sequence
c	43	14.8	46.2	38	6	AR332379	AR332379 Sequence
c	44	14.8	46.2	38	6	AR332820	AR332820 Sequence
c	45	14.8	46.2	38	6	AX222448	AX222448 Sequence

ALIGNMENTS

RESULT 1
AX453842
LOCUS AX453842 32 bp RNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent EP1213351.
ACCESSION AX453842
VERSION AX453842.1 GI:21713511
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Taira,K., Warashina,M. and Warashina,T.
TITLE Nucleic acid enzymes acquiring an activity for cleaving a target rna by recognising another molecule
JOURNAL Patent: EP 1213351-A 1 12-JUN-2002;
National Institute of Advanced Industrial Science and Technology (JP)
FEATURES Location/Qualifiers
source 1. .32
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
/note="maxizyme-constituting RNA mol ecule"
ORIGIN

Query Match 100.0%; Score 32; DB 6; Length 32;
Best Local Similarity 71.9%; Pred. No. 0.00063;
Matches 23; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGUCCUGGCCUGAUGAGAGUGAUGAGCUCUUC 32
||:||:||||:||:|||||:||:||||:|::|
Db 1 GGTCCTGGCCTGATGAGAGTGATGAGCTCTTC 32

RESULT 2
BD143498
LOCUS BD143498 32 bp RNA linear PAT 17-JAN-2003
DEFINITION Nucleic acid enzyme acquiring activity of cleaving other specific target RNA by recognizing RNA molecule.
ACCESSION BD143498
VERSION BD143498.1 GI:27849256
KEYWORDS JP 2002119283-A/1.

```
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule
  JOURNAL   Patent: JP 2002119283-A 1 23-APR-2002;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND P T BAKRIE AND BROTHERS,JAPAN BIO INDUSTRY ASSOCIATION,
            TECHNOLOGY
COMMENT     OS   Artificial Sequence
            PN   JP 2002119283-A/1
            PD   23-APR-2002
            PF   13-OCT-2000 JP 2000313320
            PI   KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO WARASHINA PC
            C12N15/09,A61K9/127,A61K38/46,A61K48/00,A61P31/12,A61P35/00, PC
            C12N1/15,
            PC   C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12Q1/25,C12Q1/68,C12N15/
            PC   00,A61K37/54,
            PC   C12N5/00
            CC   Description of Artificial Sequence: maxizyme-constituting RNA
            CC           molecule
            FH   Key             Location/Qualifiers
            FT   source          1. .32
            FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
     source         1. .32
                    /organism="synthetic construct"
                    /mol_type="genomic RNA"
                    /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 32;  DB 6;  Length 32;
  Best Local Similarity  71.9%;  Pred. No. 0.00063;
  Matches   23;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAUGAGCUCUUC 32
              ||:||:||||:||:|||||:||:||||:|::|
Db          1 GGTCCTGGCCTGATGAGAGTGATGAGCTCTTC 32


RESULT 3
AX453844
LOCUS       AX453844                  33 bp    RNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 3 from Patent EP1213351.
ACCESSION   AX453844
VERSION     AX453844.1  GI:21713513
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzymes acquiring an activity for cleaving a target
            rna by recognising another molecule
  JOURNAL   Patent: EP 1213351-A 3 12-JUN-2002;
            National Institute of Advanced Industrial Science and Technology
            (JP)
FEATURES            Location/Qualifiers
     source         1. .33
                    /organism="synthetic construct"
                    /mol_type="unassigned RNA"
                    /db_xref="taxon:32630"
                    /note="maxizyme-constituting RNA mol ecule"
ORIGIN

  Query Match           66.9%;  Score 21.4;  DB 6;  Length 33;
  Best Local Similarity  69.6%;  Pred. No. 68;
  Matches   16;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAU 23
              ||:||:||||:||:|||||: |:
Db          1 GGTCCTGGCCTGATGAGAGTTAT 23


RESULT 4
BD143500
LOCUS       BD143500                  33 bp    RNA     linear   PAT 17-JAN-2003
DEFINITION  Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule.
ACCESSION   BD143500
VERSION     BD143500.1  GI:27849258
KEYWORDS    JP 2002119283-A/3.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule
  JOURNAL   Patent: JP 2002119283-A 3 23-APR-2002;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND P T BAKRIE AND BROTHERS,JAPAN BIO INDUSTRY ASSOCIATION,
            TECHNOLOGY
COMMENT     OS   Artificial Sequence
            PN   JP 2002119283-A/3
            PD   23-APR-2002
            PF   13-OCT-2000 JP 2000313320
            PI   KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO WARASHINA PC
            C12N15/09,A61K9/127,A61K38/46,A61K48/00,A61P31/12,A61P35/00, PC
            C12N1/15,
            PC   C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12Q1/25,C12Q1/68,C12N15/
            PC   00,A61K37/54,
            PC   C12N5/00
            CC   Description of Artificial Sequence: maxizyme-constituting RNA
            CC           molecule
            FH   Key             Location/Qualifiers
            FT   source          1. .33
            FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
     source         1. .33
                    /organism="synthetic construct"
                    /mol_type="genomic RNA"
                    /db_xref="taxon:32630"
ORIGIN

  Query Match           66.9%;  Score 21.4;  DB 6;  Length 33;
  Best Local Similarity  69.6%;  Pred. No. 68;
  Matches   16;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAU 23
              ||:||:||||:||:|||||: |:
Db          1 GGTCCTGGCCTGATGAGAGTTAT 23


RESULT 5
CQ547198
LOCUS       CQ547198                  60 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 16833 from Patent WO0210449.
ACCESSION   CQ547198
VERSION     CQ547198.1  GI:41513462
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
  TITLE     Oligonucleotide library for detecting rna transcripts and splice
            variants that populate a transcriptome
  JOURNAL   Patent: WO 0210449-A 16833 07-FEB-2002;
            Compugen Inc. (US)
FEATURES            Location/Qualifiers
     source         1. .60
                    /organism="Homo sapiens"
```

```
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             55.6%;  Score 17.8;  DB 6;  Length 60;
  Best Local Similarity   71.4%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUG 21
              ||:||:||||:|| || ||:|
Db         37 GGTCCTGGCCTGAAGACAGTG 57


RESULT 6
BD197186
LOCUS       BD197186                  29 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Method and reagent for treating diseases or conditions concerning
            molecule participating in vasculogenic response.
ACCESSION   BD197186
VERSION     BD197186.1  GI:33006956
KEYWORDS    JP 2002509721-A/212.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 29)
  AUTHORS   Pavco,P.A., Roberts,E., Jarvis,T., Coeshott,C. and Mcswiggen,J.A.
  TITLE     Method and reagent for treating diseases or conditions concerning
            molecule participating in vasculogenic response
  JOURNAL   Patent: JP 2002509721-A 212 02-APR-2002;
            RIBOZYME PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002509721-A/212
            PD   02-APR-2002
            PF   24-MAR-1999 JP 2000541291
            PR   27-MAR-1998 US    60/079678
            PI   PAMELA A PAVCO,ELISABETH ROBERTS,THALE JARVIS,CLAIRE COESHOTT,
            PI   JAMES A MCSWIGGEN
            PC
            C12N15/09,A61K31/7088,A61K31/7125,A61K48/00,A61P3/10,A61P17/06, PC
             A61P29/00,
            PC   A61P35/00,A61P43/00,C12N5/10,C12N9/00//A61K35/76,C12N15/00, PC
              C12N5/00
            CC   Synthesized Hammerhead Ribozyme
            CC   The letter 'n' stands for any base or bases forming a loop or
            CC            stem-loop
            CC   that may contain multiple nucleic acid analogues or 2'- CC
                   deoxynucleotides.
            FH   Key             Location/Qualifiers
            FT   source          1. .29
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .29
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             54.4%;  Score 17.4;  DB 6;  Length 29;
  Best Local Similarity   53.6%;  Pred. No. 5.4e+03;
  Matches   15;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 UCCUGGCCUGAUGAGAGUGAUGAGCUCU 30
              :| :||||:||:|||    | ||| :|:
Db          2 TCTTGGCCTGATGAGNCGAAAGAGTTCT 29


RESULT 7
AX453853
LOCUS       AX453853                  40 bp    RNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 12 from Patent EP1213351.
ACCESSION   AX453853
VERSION     AX453853.1  GI:21713522
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzymes acquiring an activity for cleaving a target
            rna by recognising another molecule
  JOURNAL   Patent: EP 1213351-A 12 12-JUN-2002;
            National Institute of Advanced Industrial Science and Technology
            (JP)
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="wild type ribozyme"
ORIGIN

  Query Match             54.4%;  Score 17.4;  DB 6;  Length 40;
  Best Local Similarity   63.0%;  Pred. No. 5.4e+03;
  Matches   17;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAUGAGC 27
              ||:||:||||:||:|||   ||   ||
Db          1 GGTCCTGGCCTGATGAGGCCGAAAGGC 27


RESULT 8
AR291770/c
LOCUS       AR291770                  47 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 3505 from patent US 6537751.
ACCESSION   AR291770
VERSION     AR291770.1  GI:31679054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 3505 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             54.4%;  Score 17.4;  DB 6;  Length 47;
  Best Local Similarity   51.7%;  Pred. No. 5.5e+03;
  Matches   15;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GUCCUGGCCUGAUGAGAGUGAUGAGCUCU 30
              |:||  ||::||:||  |   :||||:|:
Db         32 GTCCATGCMTGATGATTGCTCTGAGCTCT 4


RESULT 9
AR179707
LOCUS       AR179707                  30 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 7 from patent US 6326482.
ACCESSION   AR179707
VERSION     AR179707.1  GI:20221262
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Stewart,T.A. and Lu,Y.
  TITLE     SH2 domain-containing peptides
  JOURNAL   Patent: US 6326482-A 7 04-DEC-2001;
FEATURES             Location/Qualifiers
     source          1. .30
```

```
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             51.9%;  Score 16.6;  DB 6;  Length 30;
  Best Local Similarity   65.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          7 GGCCUGAUGAGAGUGAUGAGCUC 29
              ||||:| :|| ||:|  ||||:|
Db          6 GGCCTGTTGAAAGTGCAGAGCTC 28


RESULT 10
BD222050
LOCUS       BD222050                  30 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  SH2 domain-containing peptide.
ACCESSION   BD222050
VERSION     BD222050.1  GI:33031820
KEYWORDS    JP 2002512032-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Stewart,T.A. and Lu,Y.
  TITLE     SH2 domain-containing peptide
  JOURNAL   Patent: JP 2002512032-A 4 23-APR-2002;
            GENENTECH INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002512032-A/4
            PD   23-APR-2002
            PF   23-APR-1999 JP 2000544799
            PR   23-APR-1998 US    60/082767,22-DEC-1998 US    60/113296 PI
            TIMOTHY A STEWART,YANMEI LU
            PC   C12N15/09,C07K14/47,C07K16/18,C07K16/42,C07K19/00,C12N1/19, PC
              C12N1/21,
            PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/02,C12Q1/68,G01N33/53, PC
            G01N33/53,
            PC   G01N33/68,C12N15/00,C12N5/00
            CC   artificial sequence
            FH   Key             Location/Qualifiers
            FT   source          1. .30
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .30
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             51.9%;  Score 16.6;  DB 6;  Length 30;
  Best Local Similarity   65.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          7 GGCCUGAUGAGAGUGAUGAGCUC 29
              ||||:| :|| ||:|  ||||:|
Db          6 GGCCTGTTGAAAGTGCAGAGCTC 28


RESULT 11
BD162072/c
LOCUS       BD162072                  27 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Method for detecting or quantitating protein.
ACCESSION   BD162072
VERSION     BD162072.1  GI:27867830
KEYWORDS    JP 2002191364-A/8.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Tanaka,M. and Kikuchi,A.
  TITLE     Method for detecting or quantitating protein
  JOURNAL   Patent: JP 2002191364-A 8 09-JUL-2002;
            MITSUBISHI CHEMICAL CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002191364-A/8
            PD   09-JUL-2002
            PF   26-DEC-2000 JP 2000394675
            PI   MASAHITO TANAKA,AKIHIKO KIKUCHI
            PC   C12N15/09,C07K14/21,C07K14/47,C07K16/40,C07K19/00,C12M1/40, PC
              C12N15/02,
            PC   C12P21/02,G01N33/53,G01N33/577//C07K16/12,C12N15/00,C12N15/00
            CC   Description of Artificial Sequence:synthesized FH   Key
                 Location/Qualifiers
            FT   source          1. .27
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .27
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             51.2%;  Score 16.4;  DB 6;  Length 27;
  Best Local Similarity   53.8%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          7 GGCCUGAUGAGAGUGAUGAGCUCUUC 32
              ||| :|| || ||:| : || :|::|
Db         27 GGCTTGAAGACAGTGGTAAGATCTTC 2


RESULT 12
AX211297
LOCUS       AX211297                  55 bp    DNA     linear   PAT 06-SEP-2001
DEFINITION  Sequence 12 from Patent WO0159137.
ACCESSION   AX211297
VERSION     AX211297.1  GI:15523706
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ulvskov,P., Shols,H., Visser,R., Borkhardt,B., Sorensen,S.O.,
            Oomen,R., Vincken,J.P., Mccain,M., Skjot,M., Bush,M., Voragen,C.D.
            and Beldman,G.
  TITLE     Method for remodelling cell wall polysaccharide structures in
            plants
  JOURNAL   Patent: WO 0159137-A 12 16-AUG-2001;
            Ulvskov, Peter (DK) ; Shols, Henk (NL) ; Visser, Richard (NL) ;
            Borkhardt, Bernhard (DK) ; Sorensen, Susanne O. (DK) ; Oomen,
            Ronald (NL) ; Vincken, Jean-Paul (NL) ; McCain, Maureen (GB) ;
            Skjot, Michael (DK) ; Bush, Max (GB) ; Voragen, Chantal Doeswijk
            (NL) ; Beldman, Gerrit (NL)
FEATURES              Location/Qualifiers
     source           1. .55
                      /organism="synthetic construct"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32630"
                      /note="ARA-KDEL primer2 used for the amplification of the
                      arabinanases encoding cDNA"
ORIGIN

  Query Match             51.2%;  Score 16.4;  DB 6;  Length 55;
  Best Local Similarity   53.8%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          5 CUGGCCUGAUGAGAGUGAUGAGCUCU 30
              |:||||:| :| ||  ||:|| |: :
Db         15 CTGGCCTGTTGTGAAGGATGAACTTT 40


RESULT 13
AR213624
LOCUS       AR213624                  39 bp    DNA     linear   PAT 25-SEP-2002
DEFINITION  Sequence 58 from patent US 6405989.
```

```
ACCESSION   AR213624
VERSION     AR213624.1  GI:23310903
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Davis,M.E., White,R.A., Saunders,C., Polin,R., Kristiansen,K.,
            Ballone,M. and Grossman,G.
  TITLE     Rollable sports base
  JOURNAL   Patent: US 6405989-A 58 18-JUN-2002;
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             50.6%;  Score 16.2;  DB 6;  Length 39;
  Best Local Similarity   62.1%;  Pred. No. 2e+04;
  Matches   18;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAUGAGCUC 29
              || || || |:| :||| |:|| || | |
Db          4 GGGCCCGGGCTGCTGAGGGTGACGATCCC 32


RESULT 14
BD057682
LOCUS       BD057682                  39 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Fusion proteins comprising bacteriophage coat protein and a
            single-chain T cell receptor.
ACCESSION   BD057682
VERSION     BD057682.1  GI:22603288
KEYWORDS    JP 2001514503-A/58.
SOURCE      Aspergillus tubingensis
  ORGANISM  Aspergillus tubingensis
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Weidanz,J.A., Card,K.F. and Wong,H.C.
  TITLE     Fusion proteins comprising bacteriophage coat protein and a
            single-chain T cell receptor
  JOURNAL   Patent: JP 2001514503-A 58 11-SEP-2001;
            SUNOL MOLECULAR CORP
COMMENT     PN   JP 2001514503-A/58
            PD   11-SEP-2001
            PF   05-MAR-1998 JP 1998537984
            PR   07-MAR-1997 US    08/813781
            PI   JON A WEIDANZ,KIMBERLYN F CARD,HING C WONG
            PC   C12Q1/68,C12N7/01,C12N15/70
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Aspergillus tubingensis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5068"
ORIGIN

  Query Match             50.6%;  Score 16.2;  DB 6;  Length 39;
  Best Local Similarity   62.1%;  Pred. No. 2e+04;
  Matches   18;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAUGAGCUC 29
              || || || |:| :||| |:|| || | |
Db          4 GGGCCCGGGCTGCTGAGGGTGACGATCCC 32


RESULT 15
BD081512
LOCUS       BD081512                  39 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Soluble single-chain T-cell receptor proteins.
ACCESSION   BD081512
VERSION     BD081512.1  GI:22627115
KEYWORDS    JP 2001519143-A/58.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Weidanz,J.A., Card,K.F. and Wong,H.C.
  TITLE     Soluble single-chain T-cell receptor proteins
  JOURNAL   Patent: JP 2001519143-A 58 23-OCT-2001;
            SUNOL MOLECULAR CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2001519143-A/58
            PD   23-OCT-2001
            PF   28-SEP-1998 JP 2000514936
            PR   02-OCT-1997 US    08/943086
            PI   JON A WEIDANZ,KIMBERLYN F CARD,HING C WONG
            PC   C12N15/09,A61K38/00,A61K39/395,A61P43/00,C07K14/725,C07K16/28,
            PC   C12P21/02//
            PC   C12P21/08,C12N15/00,A61K37/02
            CC   Description of Artificial Sequence: primer
            FH   Key             Location/Qualifiers
            FT   source          1. .39
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             50.6%;  Score 16.2;  DB 6;  Length 39;
  Best Local Similarity   62.1%;  Pred. No. 2e+04;
  Matches   18;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 GGUCCUGGCCUGAUGAGAGUGAUGAGCUC 29
              || || || |:| :||| |:|| || | |
Db          4 GGGCCCGGGCTGCTGAGGGTGACGATCCC 32


Search completed: October 31, 2004, 03:24:20
Job time : 862.61 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 01:51:44 ; Search time 719.39 Seconds
(without alignments)
1774.868 Million cell updates/sec

Title: US-09-974-974-2
Perfect score: 27
Sequence: 1 gucugacuguucaucgaaaccgggucc 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 1924030

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	27	100.0	27	6	AX453843	AX453843 Sequence
	2	27	100.0	27	6	BD143499	BD143499 Nucleic a
	3	18	66.7	29	6	AX453845	AX453845 Sequence
	4	18	66.7	29	6	BD143501	BD143501 Nucleic a
c	5	17	63.0	44	6	AR183095	AR183095 Sequence
c	6	16	59.3	49	6	CQ760641	CQ760641 Sequence
c	7	15.6	57.8	45	9	HSCD85716	Z50829 H.sapiens m
c	8	15.2	56.3	43	6	AX297640	AX297640 Sequence
	9	15	55.6	24	6	AX445798	AX445798 Sequence
c	10	15	55.6	51	6	CQ007470	CQ007470 Sequence
	11	14.8	54.8	32	6	AR090340	AR090340 Sequence
	12	14.8	54.8	32	6	AR197375	AR197375 Sequence
	13	14.8	54.8	32	6	AR259529	AR259529 Sequence
c	14	14.8	54.8	33	6	BD274953	BD274953 POLYNUCLE
c	15	14.8	54.8	33	6	AR474082	AR474082 Sequence
c	16	14.6	54.1	22	6	AR078172	AR078172 Sequence
	17	14.6	54.1	28	6	AX351504	AX351504 Sequence
c	18	14.6	54.1	54	10	AF224227	AF224227 Mus muscu
	19	14.4	53.3	58	6	AX773271	AX773271 Sequence
	20	14.2	52.6	60	6	CQ545113	CQ545113 Sequence
	21	14.2	52.6	60	12	SYN322RSA	M11126 Plasmid pBR
	22	14	51.9	60	6	CQ548059	CQ548059 Sequence
	23	13.8	51.1	20	6	BD101849	BD101849 Kidney re
	24	13.8	51.1	23	11	DOGTCRAA	L77454 Canis famil
c	25	13.8	51.1	28	6	A21821	A21821 Polynucleot
c	26	13.8	51.1	28	6	A22663	A22663 Oligonucleo
c	27	13.8	51.1	28	6	AR094600	AR094600 Sequence
c	28	13.8	51.1	28	6	AR119613	AR119613 Sequence
c	29	13.8	51.1	28	6	I21787	I21787 Sequence 1
c	30	13.8	51.1	28	6	AR218766	AR218766 Sequence
c	31	13.8	51.1	28	6	AR306217	AR306217 Sequence
c	32	13.8	51.1	28	6	AR360569	AR360569 Sequence
c	33	13.8	51.1	32	6	AR090246	AR090246 Sequence
c	34	13.8	51.1	32	6	AR197281	AR197281 Sequence
c	35	13.8	51.1	32	6	AR259435	AR259435 Sequence
c	36	13.8	51.1	48	6	CQ654561	CQ654561 Sequence
c	37	13.8	51.1	51	6	AX158927	AX158927 Sequence
c	38	13.8	51.1	56	6	A36725	A36725 Sequence 13
	39	13.8	51.1	60	6	CQ544169	CQ544169 Sequence
	40	13.8	51.1	60	6	CQ550275	CQ550275 Sequence
c	41	13.6	50.4	20	6	AR442551	AR442551 Sequence
	42	13.6	50.4	24	6	AX290913	AX290913 Sequence
	43	13.6	50.4	26	6	AR149782	AR149782 Sequence
	44	13.6	50.4	26	6	AR156748	AR156748 Sequence
	45	13.6	50.4	26	6	BD175816	BD175816 Novel tig

ALIGNMENTS

RESULT 1
AX453843
LOCUS AX453843 27 bp RNA linear PAT 06-JUL-2002
DEFINITION Sequence 2 from Patent EP1213351.
ACCESSION AX453843
VERSION AX453843.1 GI:21713512
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Taira,K., Warashina,M. and Warashina,T.
TITLE Nucleic acid enzymes acquiring an activity for cleaving a target rna by recognising another molecule
JOURNAL Patent: EP 1213351-A 2 12-JUN-2002;
National Institute of Advanced Industrial Science and Technology (JP)
FEATURES Location/Qualifiers
source 1. .27
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
/note="maxizyme-constituting RNA mol ecule"
ORIGIN

Query Match 100.0%; Score 27; DB 6; Length 27;
Best Local Similarity 74.1%; Pred. No. 0.013;
Matches 20; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUCUGACUGUUCAUCGAAACCGGGUCC 27
|:|:|||:|::||:||||||||||:||
Db 1 GTCTGACTGTTCATCGAAACCGGGTCC 27

RESULT 2
BD143499
LOCUS BD143499 27 bp RNA linear PAT 17-JAN-2003
DEFINITION Nucleic acid enzyme acquiring activity of cleaving other specific target RNA by recognizing RNA molecule.
ACCESSION BD143499
VERSION BD143499.1 GI:27849257
KEYWORDS JP 2002119283-A/2.

```
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule
  JOURNAL   Patent: JP 2002119283-A 2 23-APR-2002;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND P T BAKRIE AND BROTHERS,JAPAN BIO INDUSTRY ASSOCIATION,
            TECHNOLOGY
COMMENT     OS   Artificial Sequence
            PN   JP 2002119283-A/2
            PD   23-APR-2002
            PF   13-OCT-2000 JP 2000313320
            PI   KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO WARASHINA PC
            C12N15/09,A61K9/127,A61K38/46,A61K48/00,A61P31/12,A61P35/00, PC
            C12N1/15,
            PC   C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12Q1/25,C12Q1/68,C12N15/
            PC   00,A61K37/54,
            PC   C12N5/00
            CC   Description of Artificial Sequence: maxizyme-constituting RNA
            CC           molecule
            FH   Key             Location/Qualifiers
            FT   source          1. .27
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 27;  DB 6;  Length 27;
  Best Local Similarity   74.1%;  Pred. No. 0.013;
  Matches   20;  Conservative    7;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUCUGACUGUUCAUCGAAACCGGGUCC 27
              |:|:|||:|::||:||||||||||:||
Db          1 GTCTGACTGTTCATCGAAACCGGGTCC 27


RESULT 3
AX453845
LOCUS       AX453845                 29 bp    RNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 4 from Patent EP1213351.
ACCESSION   AX453845
VERSION     AX453845.1  GI:21713514
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzymes acquiring an activity for cleaving a target
            rna by recognising another molecule
  JOURNAL   Patent: EP 1213351-A 4 12-JUN-2002;
            National Institute of Advanced Industrial Science and Technology
            (JP)
FEATURES             Location/Qualifiers
     source          1. .29
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="maxizyme-constituting RNA mol ecule"
ORIGIN

  Query Match             66.7%;  Score 18;  DB 6;  Length 29;
  Best Local Similarity   77.8%;  Pred. No. 4.9e+02;
  Matches   14;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 UUCAUCGAAACCGGGUCC 27
              ::||:||||||||||:||
Db         12 TTCATCGAAACCGGGTCC 29


RESULT 4
BD143501
LOCUS       BD143501                 29 bp    RNA     linear   PAT 17-JAN-2003
DEFINITION  Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule.
ACCESSION   BD143501
VERSION     BD143501.1  GI:27849259
KEYWORDS    JP 2002119283-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 29)
  AUTHORS   Taira,K., Warashina,M. and Warashina,T.
  TITLE     Nucleic acid enzyme acquiring activity of cleaving other specific
            target RNA by recognizing RNA molecule
  JOURNAL   Patent: JP 2002119283-A 4 23-APR-2002;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
            SCIENCE AND P T BAKRIE AND BROTHERS,JAPAN BIO INDUSTRY ASSOCIATION,
            TECHNOLOGY
COMMENT     OS   Artificial Sequence
            PN   JP 2002119283-A/4
            PD   23-APR-2002
            PF   13-OCT-2000 JP 2000313320
            PI   KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO WARASHINA PC
            C12N15/09,A61K9/127,A61K38/46,A61K48/00,A61P31/12,A61P35/00, PC
            C12N1/15,
            PC   C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12Q1/25,C12Q1/68,C12N15/
            PC   00,A61K37/54,
            PC   C12N5/00
            CC   Description of Artificial Sequence: maxizyme-constituting RNA
            CC           molecule
            FH   Key             Location/Qualifiers
            FT   source          1. .29
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .29
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             66.7%;  Score 18;  DB 6;  Length 29;
  Best Local Similarity   77.8%;  Pred. No. 4.9e+02;
  Matches   14;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 UUCAUCGAAACCGGGUCC 27
              ::||:||||||||||:||
Db         12 TTCATCGAAACCGGGTCC 29


RESULT 5
AR183095/c
LOCUS       AR183095                 44 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 72 from patent US 6340461.
ACCESSION   AR183095
VERSION     AR183095.1  GI:20226688
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Terman,D.Stephen.
  TITLE     Superantigen based methods and compositions for treatment of
            diseases
  JOURNAL   Patent: US 6340461-A 72 22-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
  Query Match             63.0%;  Score 17;  DB 6;  Length 44;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 CUGACUGUUCAUCGAAACCGGGUCC 27
              | |||:|::||: | | |||| :||
Db         25 CAGACTGTTCATAGGATCCGGATCC 1


RESULT 6
CQ760641/c
LOCUS       CQ760641                  49 bp    DNA     linear   PAT 03-MAR-2004
DEFINITION  Sequence 83 from Patent WO2004003229.
ACCESSION   CQ760641
VERSION     CQ760641.1  GI:44904144
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Nex,B.R., Vogel,U., Rockenbauer,E. and Bukowy,Z.K.
  TITLE     Disease risk estimating method using sequence polymorphisms in a
            specific region of chromosome 19
  JOURNAL   Patent: WO 2004003229-A 83 08-JAN-2004;
            Aarhus University (DK) ; Arbejdsmilj Instituttet (National
            Institute of Occupational Health) (DK)
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Probe"
ORIGIN

  Query Match             59.3%;  Score 16;  DB 6;  Length 49;
  Best Local Similarity   54.2%;  Pred. No. 5.5e+03;
  Matches   13;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GUCUGACUGUUCAUCGAAACCGGG 24
              |:|:|| : :: |:|||| || ||
Db         28 GTCTGAGTATTGATCGAACCCAGG 5


RESULT 7
HSCD85716/c
LOCUS       HSCD85716                 45 bp    mRNA    linear   PRI 15-MAR-1996
DEFINITION  H.sapiens mRNA for T cell receptor (ID CD85716).
ACCESSION   Z50829
VERSION     Z50829.1  GI:1004252
KEYWORDS    constant region; joining region; T cell receptor; variable region.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Wang,E.C., Moss,P.A., Frodsham,P., Lehner,P.J., Bell,J.I. and
            Borysiewicz,L.K.
  TITLE     CD8highCD57+ T lymphocytes in normal, healthy individuals are
            oligoclonal and respond to human cytomegalovirus
  JOURNAL   J. Immunol. 155 (10), 5046-5056 (1995)
  MEDLINE   96062329
   PUBMED   7594513
REFERENCE   2  (bases 1 to 45)
  AUTHORS   Wang,E.C., Moss,P.A., Frodsham,P., Lehner,P.J., Bell,J.I. and
            Borysiewicz,L.K.
  TITLE     Oligoclonal CD8+, CD57+ T cells in normal, healthy individuals and
            their responses to HCMV
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 45)
  AUTHORS   Wang,E.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1995) Eddie CY Wang, Medicine, University of
            Wales of College of, Medicine, Heath Park, Cardiff, South
            Glamorgan, CF4 4XN, Wales UK
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /isolate="SG"
                     /db_xref="taxon:9606"
     V_segment       1. .12
                     /note="end T cell receptor V beta 3"
     C_region        13. .27
                     /note="CDR3"
     J_segment       28. .45
                     /note="beginning J beta 2.1"
ORIGIN

  Query Match             57.8%;  Score 15.6;  DB 9;  Length 45;
  Best Local Similarity   63.6%;  Pred. No. 8.7e+03;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 ACUGUUCAUCGAAACCGGGUCC 27
              ||:| :||: | ||||| |:||
Db         40 ACTGCTCATTGTAACCGAGTCC 19


RESULT 8
AX297640/c
LOCUS       AX297640                  43 bp    DNA     linear   PAT 21-NOV-2001
DEFINITION  Sequence 9402 from Patent WO0179548.
ACCESSION   AX297640
VERSION     AX297640.1  GI:17059331
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Barany,F., Zirvi,M., Gerry,N.P., Favis,R. and Kliman,R.
  TITLE     Method of designing addressable array for detection of nucleic acid
            sequence differences using ligase detection reaction
  JOURNAL   Patent: WO 0179548-A 9402 25-OCT-2001;
            CORNELL RESEARCH FOUNDATION, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Hypothetical Probe Sequence"
ORIGIN

  Query Match             56.3%;  Score 15.2;  DB 6;  Length 43;
  Best Local Similarity   65.0%;  Pred. No. 1.4e+04;
  Matches   13;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 CUGUUCAUCGAAACCGGGUC 26
              |:|: ||:|||| |||||:|
Db         35 CTGTCCATCGAATGCGGGTC 16


RESULT 9
AX445798
LOCUS       AX445798                  24 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 2253 from Patent WO0216649.
ACCESSION   AX445798
VERSION     AX445798.1  GI:21694697
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Gunderson,K.
  TITLE     Probes and decoder oligonucleotides
  JOURNAL   Patent: WO 0216649-A 2253 28-FEB-2002;
            Illumina, Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Computer Generated Probe Sequence."
ORIGIN

  Query Match             55.6%;  Score 15;  DB 6;  Length 24;
  Best Local Similarity   60.9%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 CUGACUGUUCAUCGAAACCGGGU 25
              |:|||:|    :|||| ||||:
Db          1 CTGACTGCATCTCGAAAGCGGGT 23


RESULT 10
CQ007470/c
LOCUS       CQ007470                  51 bp    DNA     linear   PAT 16-JAN-2004
DEFINITION  Sequence 6110 from Patent WO0147944.
ACCESSION   CQ007470
VERSION     CQ007470.1  GI:41014119
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0147944-A 6110 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Accession number cg43988954"
ORIGIN

  Query Match             55.6%;  Score 15;  DB 6;  Length 51;
  Best Local Similarity   60.9%;  Pred. No. 1.8e+04;
  Matches   14;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GUCUGACUGUUCAUCGAAACCGG 23
              | |:|  :| :| :|||||||||
Db         23 GCCTGGATGGTCTTCGAAACCGG 1


RESULT 11
AR090340
LOCUS       AR090340                  32 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 460 from patent US 5994076.
ACCESSION   AR090340
VERSION     AR090340.1  GI:10017095
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 5994076-A 460 30-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             54.8%;  Score 14.8;  DB 6;  Length 32;
  Best Local Similarity   53.8%;  Pred. No. 2.1e+04;
  Matches   14;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUCC 27
              :| || : : ||:  ||||| ||:||
Db          6 TCAGAATCTCCATTCAAACCAGGTCC 31


RESULT 12
AR197375
LOCUS       AR197375                  32 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 460 from patent US 6352829.
ACCESSION   AR197375
VERSION     AR197375.1  GI:20247224
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 6352829-A 460 05-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             54.8%;  Score 14.8;  DB 6;  Length 32;
  Best Local Similarity   53.8%;  Pred. No. 2.1e+04;
  Matches   14;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUCC 27
              :| || : : ||:  ||||| ||:||
Db          6 TCAGAATCTCCATTCAAACCAGGTCC 31


RESULT 13
AR259529
LOCUS       AR259529                  32 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 460 from patent US 6489455.
ACCESSION   AR259529
VERSION     AR259529.1  GI:27310040
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 6489455-A 460 03-DEC-2002;
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             54.8%;  Score 14.8;  DB 6;  Length 32;
  Best Local Similarity   53.8%;  Pred. No. 2.1e+04;
  Matches   14;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUCC 27
              :| || : : ||:  ||||| ||:||
Db          6 TCAGAATCTCCATTCAAACCAGGTCC 31


RESULT 14
BD274953/c
LOCUS       BD274953                  33 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY.
ACCESSION   BD274953
VERSION     BD274953.1  GI:33084721
KEYWORDS    JP 2002538786-A/17.
SOURCE      synthetic construct
```

```
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Shimkets,R.A.
  TITLE     POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
  JOURNAL   Patent: JP 2002538786-A 17 19-NOV-2002;
            CuraGen Corporation,Richard A Shimkets
COMMENT     OS   Artificial Sequence
            PN   JP 2002538786-A/17
            PD   19-NOV-2002
            PF   09-MAR-2000 JP 2000603363
            PR   08-MAR-2000 US    09/520781,09-MAR-1999 US    60/123667 PI
            richard a shimkets
            CC   Description of Artificial Sequence: Primer
            FH   Key             Location/Qualifiers.
FEATURES              Location/Qualifiers
     source           1. .33
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             54.8%;  Score 14.8;  DB 6;  Length 33;
  Best Local Similarity   53.8%;  Pred. No. 2.1e+04;
  Matches   14;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUCC 27
              :| || : ::|   |||||||| :||
Db         26 TCAGAATCTTCTGGGAAACCGGATCC 1


RESULT 15
AR474082/c
LOCUS       AR474082                 33 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 33 from patent US 6689866.
ACCESSION   AR474082
VERSION     AR474082.1  GI:42712839
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Shimkets,R.A.
  TITLE     Polynucleotides and proteins encoded thereby
  JOURNAL   Patent: US 6689866-A 33 10-FEB-2004;
FEATURES              Location/Qualifiers
     source           1. .33
                      /organism="unknown"
                      /mol_type="genomic DNA"
ORIGIN

  Query Match             54.8%;  Score 14.8;  DB 6;  Length 33;
  Best Local Similarity   53.8%;  Pred. No. 2.1e+04;
  Matches   14;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 UCUGACUGUUCAUCGAAACCGGGUCC 27
              :| || : ::|   |||||||| :||
Db         26 TCAGAATCTTCTGGGAAACCGGATCC 1


Search completed: October 31, 2004, 03:24:23
Job time : 722.39 secs
```